Title

Understanding atmospheric organic aerosols via factor analysis of aerosol mass spectrometry: a review.

Journal

Analytical and bioanalytical chemistry, 401(10)

ISSN

1618-2642

Authors

Zhang, Qi
Jimenez, Jose L
Canagaratna, Manjula R
et al.

Publication Date

2011-12-01

DOI

10.1007/s00216-011-5355-y

Peer reviewed

Anal Bioanal Chem (2011) 401:3045–3067
DOI 10.1007/s00216-011-5355-y

REVIEW

Understanding atmospheric organic aerosols via factor analysis of aerosol mass spectrometry: a review

Qi Zhang · Jose L. Jimenez · Manjula R. Canagaratna · Ingrid M. Ulbrich · Nga L. Ng · Douglas R. Worsnop · Yele Sun

Received: 16 June 2011 / Revised: 21 August 2011 / Accepted: 22 August 2011 / Published online: 5 October 2011

Abstract Organic species are an important but poorly characterized constituent of airborne particulate matter. A quantitative understanding of the organic fraction of particles (organic aerosol, OA) is necessary to reduce some of the largest uncertainties that confound the assessment of the radiative forcing of climate and air quality management policies. In recent years, aerosol mass spectrometry has been increasingly relied upon for highly time-resolved characterization of OA chemistry and for elucidation of aerosol sources and lifecycle processes. Aerodyne aerosol mass spectrometers (AMS) are particularly widely used, because of their ability to quantitatively characterize the size-resolved composition of submicron particles (PM_1). AMS report the bulk composition and temporal variations of OA in the form of ensemble mass spectra (MS) acquired over short time intervals. Because each MS represents the linear superposition of the spectra of individual components weighed by their concentrations, multivariate factor analysis of the MS matrix has proved effective at retrieving OA factors that offer a quantitative and simplified description of the thousands of individual organic species. The sum of the factors accounts for nearly 100% of the OA mass and each individual factor typically corresponds to a large group of OA constituents with similar chemical composition and temporal behavior that are characteristic of different sources and/or atmospheric processes. The application of this technique in aerosol mass spectrometry has grown rapidly in the last six years. Here we review multivariate factor analysis techniques applied to AMS and other aerosol mass spectrometers, and summarize key findings from field observations. Results that provide valuable information about aerosol sources and, in particular, secondary OA evolution on regional and global scales are highlighted. Advanced methods, for example a-priori constraints on factor mass spectra and the application of factor analysis to combined aerosol and gas phase data are discussed. Integrated analysis of worldwide OA factors is used to present a holistic regional and global description of OA. Finally, different ways in which OA factors can constrain global and regional models are discussed.

Keywords Aerosol mass spectrometer (AMS) · Positive matrix factorization (PMF) · Bilinear factorization · Primary

Published in the special issue *Aerosol Analysis* with Guest Editor Ralf Zimmermann.

Q. Zhang (✉) · Y. Sun
Department of Environmental Toxicology,
University of California,
1 Shields Ave,
Davis, CA 95616, USA
e-mail: dkwzhang@ucdavis.edu

J. L. Jimenez · I. M. Ulbrich
Department of Chemistry and Biochemistry,
University of Colorado at Boulder,
Boulder, CO 80309, USA

J. L. Jimenez
Cooperative Institute for Research in Environmental Sciences,
University of Colorado at Boulder,
Boulder, CO 80309, USA

M. R. Canagaratna · N. L. Ng · D. R. Worsnop
Aerodyne Research, Inc,
Billerica, MA 01821-3976, USA

Present Address:
N. L. Ng
Georgia Institute of Technology,
Atlanta, GA 30332, USA

Present Address:
Y. Sun
Institute of Atmospheric Physics, Chinese Academy of Science,
Beijing 100029, China

organic aerosol (POA) · Secondary organic aerosol (SOA) · Aerosol sources · Aerosol processes · Aerosol life cycle

Abbreviations

AMS	Aerosol mass spectrometer
BBOA	Biomass-burning organic aerosol
CCN	Cloud condensation nuclei
CMB	Chemical mass balance
COA	Cooking-related organic aerosol
CPCA	Custom principal-component analysis
EI	Electron-impact ionization
FA-AMS	Factor analysis of aerosol mass spectrometry
HOA	Hydrocarbon-like organic aerosol
HR	High resolution
HR-ToF-AMS	Aerodyne aerosol mass spectrometer equipped with a high-resolution time-of-flight mass spectrometer. It provides mass spectra with mass resolution of up to ~5000–6000
IVOC	Intermediate volatility organic compounds
LV-OOA	Low-volatility oxygenated organic aerosol
LS-AMS	Light-scattering aerosol mass spectrometer
MCA	Multiple component analysis
ME	Multilinear engine
MS	Mass spectra
NOA	Nitrogen-enriched organic aerosol
NR	Non-refractory
OA	Organic aerosol, i.e., the organic fraction of airborne particulate matter
OOA	Oxygenated organic aerosol
PAH	Polycyclic aromatic hydrocarbons
PIAMS	Photoionization aerosol mass spectrometer
PMF	Positive matrix factorization
POA	Primary organic aerosol (emitted into the atmosphere directly in the particle phase)
PTR-MS	Proton-transfer-reaction mass spectrometer
Q-AMS	Quadrupole-based aerodyne aerosol mass spectrometer. It provides unit-resolution mass spectra
SI-AMS	Soft-ionization aerosol mass spectrometers
SMPS	Scanning mobility particle sizer
SOA	Secondary organic aerosol (created by chemical reactions leading to a decrease in species volatility and increased partitioning to the particle phase)
SVOC	Semi-volatile organic compounds
SV-OOA	Semi-volatile oxygenated organic aerosol
TAG	Thermal desorption aerosol gas chromatography–mass spectrometry–flame ionization detection
TD	Thermodenuder
UMR	Unit-mass resolution
VOC	Volatile organic compounds
WSOC	Water-soluble organic carbon

Introduction

Aerosol particles affect the radiative budget of the Earth's atmosphere through scattering and absorption of light (i.e., direct climate forcing effect) and by modulating the formation and properties of clouds (i.e., indirect climate forcing effect) [1, 2]. Aerosols also have serious adverse effects on air quality [3], human health [4], and ecosystems [5]. Organic aerosol (OA, i.e., the organic fraction of particles) accounts for a substantial fraction (~10–90% [6–8]) of the global submicron aerosol burden and thus is a key determinant of aerosol properties and effects. A thorough understanding of the characteristics, sources, and processes of OA is necessary to address aerosol-related environmental issues and to improve the predictive capability of air quality and climate models.

The characterization of OA chemical composition and mass concentration is limited by analytical challenges arising from the fact that atmospheric OA comprises thousands of compounds with vastly different properties such as oxidation state, volatility, and hygroscopicity [7–9]. This compositional complexity of OA is a consequence of the extremely diverse sources and reactions of organic species in the atmosphere [10]. By a broad classification of sources, there are primary OA (POA) emitted directly in particulate form, e.g., from fossil fuel and biomass burning or mechanical processes, and secondary OA (SOA) produced from the oxidation of volatile organic compounds (VOCs) [7]. POA and VOCs are released from various biogenic, biomass burning, and anthropogenic sources [10]; SOA formation occurs via many reaction pathways that convert VOCs into low volatility species [11]. Furthermore, the composition and properties of both POA and SOA may change dynamically throughout aerosol lifetime, because of intertwined processes including emission, oxidation, fragmentation, oligomerization, gas-to-particle partitioning, and cloud processing [12–18].

Factor analysis of time and compositionally-resolved OA data enables the extraction of broad "factors" or "compo-

nents" representing species that correlate in time. Each factor extracted in this way typically corresponds to many individual molecules and contains information about their sources, processing histories, and/or chemical properties. Several publications have reported factor analysis of speciated OA data from filters [19–22], albeit with low time resolution (typically 24 h) which obscures some of the contrast in concentration variations because of the dynamics of the sources, chemistry, and transport. In recent years, online aerosol mass spectrometers have enabled chemical analyses of aerosols in real time with high time resolution (seconds to minutes) [23]. A range of mass spectrometers using various particle vaporization and ionization techniques have been developed. The most common designs include thermal desorption followed by electron ionization [24, 25] or other types of ionization [26, 27] and laser ablation [28, 29]. Instruments based on thermal desorption are mainly configured for determining ensemble particle properties averaged over defined time periods, whereas the laser-based instruments are primarily used for single-particle measurements. Single-step laser ablation and ionization detection schemes typically provide qualitative information about aerosol chemical constituents because the observed intensities are highly affected by the actual composition of the detected aerosol particle. Two-step thermal desorption and linear ionization schemes, on the other hand, provide quantitative and linearly additive mass spectra of mixtures (i.e., each mass spectrum that is observed is a linear combination of the responses from individual compounds present in the mixture). As discussed in more detail in the section "Bilinear modeling", the factor analysis methods discussed in this review are based on linear additivity of constant factor mass spectra. Thus, we focus only on analysis of data from the thermal vaporization-based aerosol mass spectrometers.

Aerodyne Research aerosol mass spectrometers (termed "AMS" hereafter) are the most widely used thermal desorption-based mass spectrometers in aerosol research. The AMS can quantify the mass concentrations of non-refractory (NR) species including sulfate, nitrate, ammonium, chloride, and total organic matter via thermal vaporization (typically at 600 °C) and 70-eV electron-impact ionization (EI) [25]. The distributions of these species as a function of particle size are also determined on the basis of measurement of particle velocities inside a vacuum chamber [30]. The HR-ToF-AMS, i.e., AMS built with a high-resolution time-of-flight mass spectrometer, is further able to determine the elemental composition and oxidation states of organic aerosols [31–33]. From each measurement, the AMS outputs an ensemble MS of OA that is the linear superposition of the mass spectra of individual species weighed by their concentrations. Because most molecules undergo extensive fragmentation during high-temperature vaporization and high-energy ionization inside the AMS, the AMS spectra provide information on the bulk composition of OA with limited molecular detail [25]. Mass spectral fragmentation can be limited by using soft-ionization aerosol mass spectrometers (SI-AMS), which afford increased information about the molecular composition of OA, although at the expense of quantifying the total OA mass, and typically of lower signal-to-noise ratio also [34–37]. The quantitative mass spectra generated by SI-AMS systems can be used in factor analysis and can provide more detailed information about OA sources.

This review summarizes the methods of factor analysis of fast time-resolved linearly additive mass spectral data, and key results obtained with these methods about primary sources, secondary formation, evolution/aging processes of atmospheric OA, and the global context of OA. One-hundred and twenty-five papers have been published in this area so far, all but three of which focus on Aerodyne AMS data; therefore the review focuses more strongly on AMS results.

Multivariate factor analysis of aerosol mass spectra

An atmospheric field study usually lasts a few weeks to months, during which an aerosol mass spectrometer operates continuously to record the temporal variations of the composition and concentration of the OA in the form of a mass spectral matrix (denoted "**ORG**" hereafter), i.e., an array of m measured mass spectral ion intensities compiled over t sampling time steps. An **ORG** from a typical AMS study, for example, usually comprises thousands of ensemble spectra acquired with a time resolution of seconds to minutes.

Bilinear modeling

The objective of multivariate factor analysis is to deconvolve the observed **ORG** matrix into unique factors (Eq. 1). Factor analysis of the data matrices from quantitative instruments (e.g., AMS and some soft-ionization aerosol mass spectrometers) usually involves solving a two-dimensional bilinear model that expresses mass conservation, such that:

$$org_{ij} = \sum_{p=1}^{P} ts_{ip} ms_{pj} + e_{ij} \quad (1)$$

where i and j refer to row and column indices, respectively, in the **ORG** matrix, org_{ij} is the signal of ion fragment j at time

step i (for the AMS it is the organic-equivalent mass concentration of that fragment, in μg m^{-3}), ts_{ip} is the concentration of a given factor p at time step i, ms_{pj} is the fractional contribution of ion fragment j in the mass spectrum of factor p, and e_{ij} is the residual not fit by the model for ion fragment j at time step i. P is the total number of factors in the solution. A graphical schematic diagram of the model is shown in Fig. 1.

Written in matrix form, Eq. (1) and Fig. 1 show that the bilinear model represents the matrix of data points (i.e., **ORG**, dimensions $t \times m$) as the product of two smaller matrices—one of which comprises the concentration time series (**TS**) and the other the mass spectra (MS) or source profiles of OA factors (total number = P)—plus a matrix of residuals (**E**) to account for the unexplained part of **ORG**:

$$\mathbf{ORG} = \mathbf{TS} \times \mathbf{MS} + \mathbf{E} \tag{2}$$

$$\mathbf{TS} = [\mathbf{ts_1}, \mathbf{ts_2}, \ldots, \mathbf{ts_P}] \tag{3}$$

$$\mathbf{MS} = \begin{bmatrix} \mathbf{ms}_1 \\ \mathbf{ms}_2 \\ \vdots \\ \mathbf{ms}_p \end{bmatrix} \tag{4}$$

As illustrated in Fig. 1, $\mathbf{ts_p}$ is a column vector representing the time series of any given factor "p" and $\mathbf{ms}_p$ is a row vector representing its mass spectrum. For the AMS, each ms is normalized to sum to 1 so that all elements in ts have units of mass concentration (μg m^{-3}). An underlying assumption of bilinear modeling is that each factor has a constant mass spectrum but varying concentration over time. If the true factor spectra are not constant as assumed by the model, **E** may be significantly larger than measurement errors even after all physically meaningful factors have been extracted.

Zhang et al. [40, 41] conducted the first bilinear factor analysis of the AMS data using a custom principal component analysis (CPCA) method. CPCA solves Eq. (1) on the basis of an iterative linear-decomposition algorithm that is initialized with the time series of two AMS tracer ions —m/z 44 (mainly CO_2^+) and m/z 57 (with a major contribution of $C_4H_9^+$ in urban areas)—as the first-guess of **TS**. Especially at urban locations, this algorithm is able to deconvolve two chemically and physically meaningful OA factors—an oxygenated OA factor (OOA) that represents SOA and a hydrocarbon-like OA factor (HOA) that represents POA associated with urban emissions [40–43]. An expanded version of the CPCA called multiple component analysis (MCA) was later developed to separate more than two factors [44]. Application of MCA to 37 AMS datasets acquired from various urban, rural, and remote atmospheric environments revealed that the sum of OOAs is often larger than the sum of HOA and other POA factors [6, 45–47], indicating that atmospheric OA are dominated by oxygenated species, mainly of secondary origin [6].

Positive matrix factorization (PMF)

PMF [48, 49] is a standard multivariate factor analysis model broadly used in the field of air pollution source

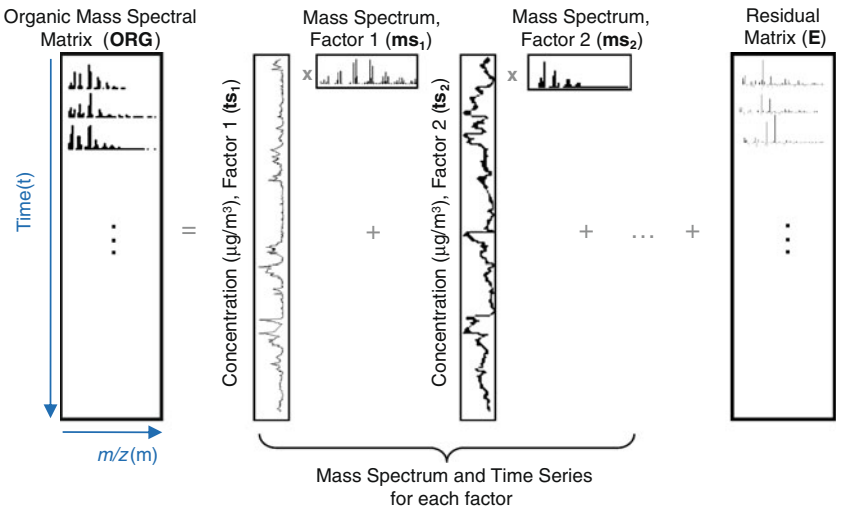

Fig. 1 Schematic diagram of bilinear factor analysis of a mass spectral matrix of an organic aerosol (**ORG**). The time series of the factors ($\mathbf{ts}_n$) make up the matrix **TS** (Eq. 3) and the mass spectra of the factors ($\mathbf{ms}_n$) make up the matrix **MS** (Eq. 4). The differences between the measurements and the modeled results are represented as the residual matrix **E**. (Adapted from Ref. [38]). An example of the factor results obtained from PMF analysis of an ambient AMS dataset is shown in Fig. 2

apportionment. In recent years, it has seen more applications in factor analysis of quantitative aerosol mass spectrometry [34, 38, 50–52]. PMF models the data matrix (**ORG**) according to Eq. (1) as a positively constrained, weighted least-squares problem without a-priori assumptions for either source (**MS**) or time (**TS**) profiles [38, 48]. The researcher chooses the number of factors, *P*, and the solution to PMF is the one that minimizes the sum of the weighed squared residuals ("*Q* value", or "PMF quality-of-fit parameter"):

$$Q = \sum_{i=1}^{t} \sum_{j=1}^{m} (e_{ij}/\sigma_{ij})^2 \quad (5)$$

where σ_{ij} is an element in the $t \times m$ matrix of estimated errors (1σ measurement precisions) corresponding to the variables (org_{ij}) in **ORG** (Eq. 1). The purpose of this scaling is to weigh each variable by its degree of measurement uncertainty, so that the factor analysis model can make use of the real information content of the dataset [48]. Each value in the solution matrices (i.e., **MS** and **TS** in Eq. 1) of PMF is constrained to be positive, reflecting the real atmospheric situation. The bilinear PMF model can be solved by several algorithms, with the PMF2 and multilinear engine (ME-2) software distributed by P. Paatero [48, 53] being the most commonly used.

If the assumptions of the bilinear model are appropriate for the dataset and the error estimates are accurate, when the minimum *Q* value is achieved all elements in the matrix are fit to within their expected error, i.e., $|e_{ij}|/\sigma_{ij} \approx 1$. Then the expected value of *Q* (Q_{exp}) should equal the degrees of freedom of the fitted data [54]:

$$Q_{exp} = (t \times m) - P \times (t + m) \quad (6)$$

For AMS datasets, because $t \times m >> P \times (t + m)$, $Q_{exp} \approx t \times m$ (the number of points in **ORG**).

Thus, if the bilinear model is appropriate and the errors are small, the solution with the correct number of factors should give Q/Q_{exp} near unity [38]. Values of $Q/Q_{exp} >> 1$ indicate either underestimation of the errors, or variability in the factor mass spectra that cannot be simply modeled as the sum of the given number of factors. $Q/Q_{exp} << 1$ indicates overestimation of the errors of the input data.

PMF analysis of aerosol mass spectra

Lanz et al. [51] reported the first PMF study on an AMS dataset acquired in Zurich, Switzerland, in summer 2005 and identified six factors, including an HOA, two OOAs, and three factors linked to charbroiling, wood burning, and food cooking sources, respectively. The less oxidized OOA-2 factor was found to represent less processed, more volatile SOA [51]. Ulbrich et al. [38] also reported using PMF for identification of a semivolatile OOA-2 factor, in addition to a more oxidized, regional OOA-1 and an HOA in Pittsburgh. These PMF results agree well with the original two-factor (HOA and OOA) CPCA results of Zhang et al. [40, 55]. The fact that PMF is able to retrieve a low-concentration, yet distinct factor (OOA-2) highlights its strength in extracting information from datasets, e.g., resolving factors that make up a small fraction of the total mass. Based on PMF analysis of synthetic **ORG** matrices that were reconstructed assuming variable contributions from Pittsburgh OOA-1, OOA-2, and HOA factors, Ulbrich et al. [38] estimated that PMF of quadrupole AMS data (unit-mass resolution; UMR) can typically retrieve factors that account for at least 5% of the AMS mass.

The extraction of two distinct OOA subfactors was achieved later in PMF studies of a large number of other AMS datasets [13, 14, 18, 56]. Typically, the more oxidized OOA factor (OOA-1) correlates well with sulfate and is thought to be more aged and non-volatile. In contrast, the less oxidized OOA (OOA-2) is thought to be typically semivolatile because of its diurnal cycles and time trends that are similar to those of ammonium nitrate and chloride, both of which dynamically partition between gaseous and particulate phases depending on ambient temperature and humidity. The relative volatility characteristics of the two OOAs were confirmed by thermodenuder measurements [57–60]. In particular, Cappa and Jimenez [60] reported volatility distributions for both OOAs and other OA components for Mexico City. For these reasons, Jimenez et al. [13] introduced the more descriptive acronyms LV-OOA (low-volatility) and SV-OOA (semivolatile), respectively, that have become the standard terminology. However, the terminology LO-OOA and MO-OOA for less and more oxidized OOA, respectively, is also appropriate, especially for datasets for which volatility data are not available. More discussion on the differences among OOA subtypes is given in the section "OOA subtypes and interpretation".

PMF studies have been conducted on AMS datasets acquired with both UMR and high-resolution (HR) mass spectrometers. Most of the earlier datasets are UMR but more HR-AMS-PMF results have been reported recently [22, 39, 50, 61–68]. A main advantage of the HR-AMS data is the separate quantification of different ions having the same nominal mass, enabling more precise characterization of the temporal variations of different ion types (e.g., $C_xH_y^+$, $C_xH_yO_z^+$, $C_xH_yN_p^+$, and $C_xH_yN_pO_z^+$). The enhanced chemical resolution, and thus the higher information content in the HR-AMS datasets, is useful for constraining the PMF solutions, reducing their rotational ambiguity and leading to more easily interpretable solutions and, potentially, a larger number of interpretable OA factors. For example Aiken et al. [64] reported that

HOA and biomass burning OA (BBOA) were better separated using HR-AMS data as opposed to when the same data were analyzed as UMR, because their spectra are somewhat similar in UMR but very different in HR. In addition, the HR mass spectra of the OA factors also contain more information useful for interpreting their sources and processing.

Figure 2 shows an example of typical PMF results from an HR-AMS dataset, including the time series, HR spectra, and diurnal patterns of the individual OA factors. The dataset was acquired in New York City in summer 2009 [39]. Five OA factors were determined, each with distinct temporal variation and mass spectral patterns:

1. LV-OOA (oxygen-to-carbon atomic ratio O/C = 0.63) that correlates strongly with sulfate;
2. SV-OOA (O/C = 0.38) that correlates better with ammonium nitrate and chloride than LV-OOA does;
3. a nitrogen-enriched OA (NOA) with a much higher N/C ratio (0.052) than other OA components (~0.004–0.011);
4. a cooking-related OA (COA) which has spectral features similar to those of POA from cooking emissions and a distinctive diurnal pattern peaking during lunch and dinner times; and
5. an HOA that represents POA from fossil fuel combustion given its low O/C ratio (0.06) and good correlation with primary combustion emission species, for example NO_x and EC.

Detailed discussions of the association of each component with different sources and processes are given in Sun et al. [39].

In addition to AMS data, PMF has been applied to the OA data from other aerosol mass spectrometers. To our knowledge, there are two published studies, one from a SI-

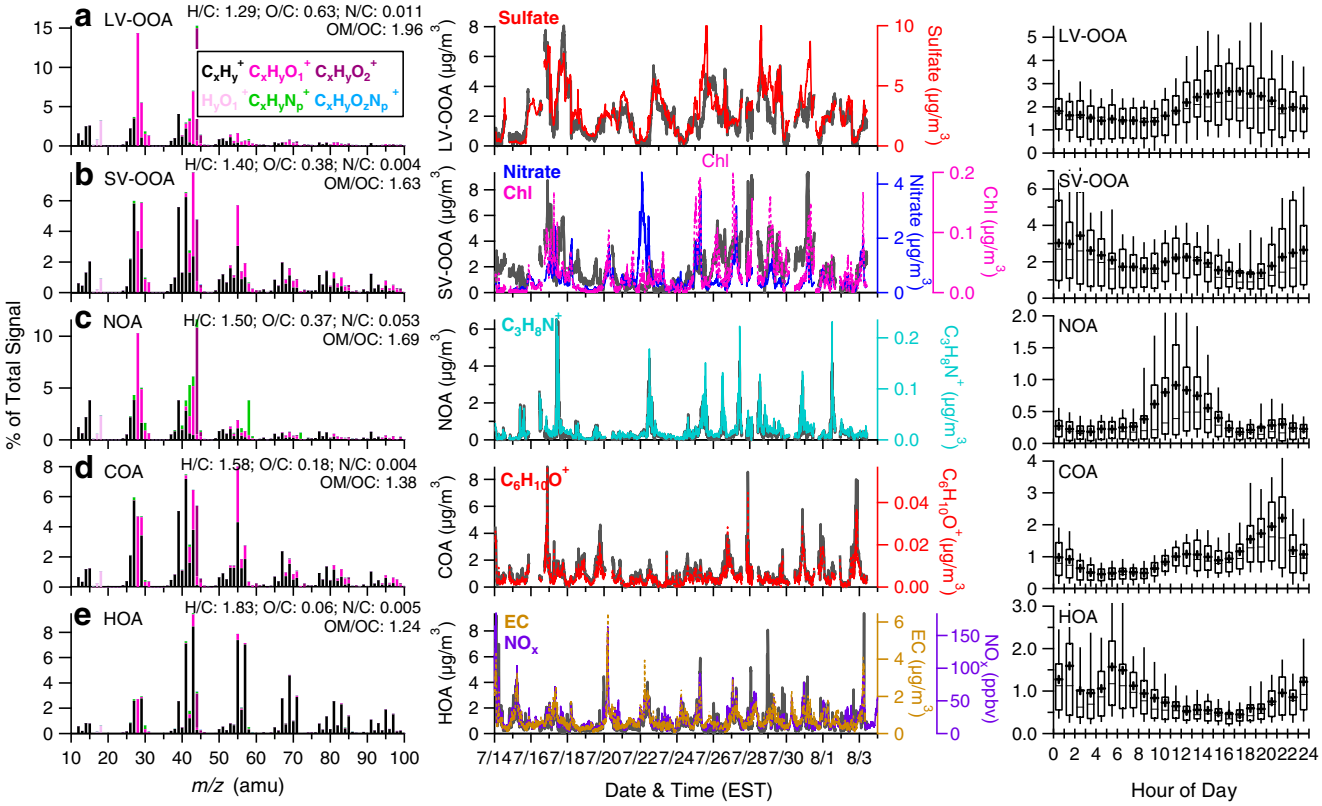

Fig. 2 The mass spectra (*left*), time series (*middle*), and diurnal patterns (*right*) of five OA factors determined on the basis of PMF analysis of an HR-ToF-AMS dataset acquired in summer 2009 in New York, NY, USA. (**a**) *LV-OOA*, surrogate for regional, highly aged, low-volatility SOA; (**b**) *SV-OOA*, surrogate for less photochemically aged, semi-volatile SOA; (**c**) *NOA*, a nitrogen-enriched OA, probably derived from an SOA formed via acid–base chemistry or photochemical reactions of amino compounds, but possibly also by other mechanisms or by involvement of other reduced nitrogen compounds; (**d**) *COA*, a POA component probably dominated by cooking emissions; and (**e**) *HOA*, a surrogate for urban, combustion-related POA. In the time series plots, the corresponding time trends of tracer compounds are: (**a**) sulfate representing low volatility secondary aerosol species formed on regional scale; (**b**) nitrate and chloride representing semivolatile secondary species; (**c**) $C_3H_8N^+$ as a tracer ion for reduced nitrogen compounds; (**d**) $C_6H_{10}O^+$ as a tracer ion for cooking aerosols; and (**e**) elemental carbon and NO_x as tracer species for combustion emissions. In the HRMS of OA factors, each peak is colored on the basis of the contributions of five ion categories: $C_xH_y^+$, $H_yO_1^+$, $C_xH_yO_z^+$, $C_xH_yN_p^+$, and $C_xH_yO_zN_p^+$. The elemental and organic mass-to-carbon ratios for each factor are shown in the legends. (Adapted from Ref. [39]) A summary of the key diagnostic plots of the PMF results is shown in Fig. 5

AMS and the other from an on-line GC–MS. Because both techniques determine individual molecules (or larger fragments or decomposition products) in aerosols, these analyses may be particularly useful for studying the sources and source contributions of OA. Williams et al. [69] performed PMF analysis on hourly time-resolution data of organic marker compounds measured with a thermal desorption aerosol GC–MS–FID (TAG) instrument from a study site in Riverside, California. The grouping of marker compounds in each factor was used to identify the presence of several different source types, including local vehicle emissions, food cooking operations, biomass burning, regional primary anthropogenic emissions, biogenic POA sources, several types of SOA, and semivolatile anthropogenic and biogenic OA. A key challenge in this type of tracer-based apportionment is to assign a fraction of OA mass to each factor, because the tracers account for only a small fraction (<20%) of the OA mass. Williams et al. [69] performed this step using a multivariate fit of their OA components to the AMS OA concentration. Figure 3 shows the diurnal cycles of OA sources obtained with this method during the summer study, with SOA (POA) being more important during the day (evening/night), in agreement with, e.g., the NYC results in Fig. 2 and AMS results at other locations. An earlier study applied principal-component analysis (PCA) on data from the same instrument, and identified several

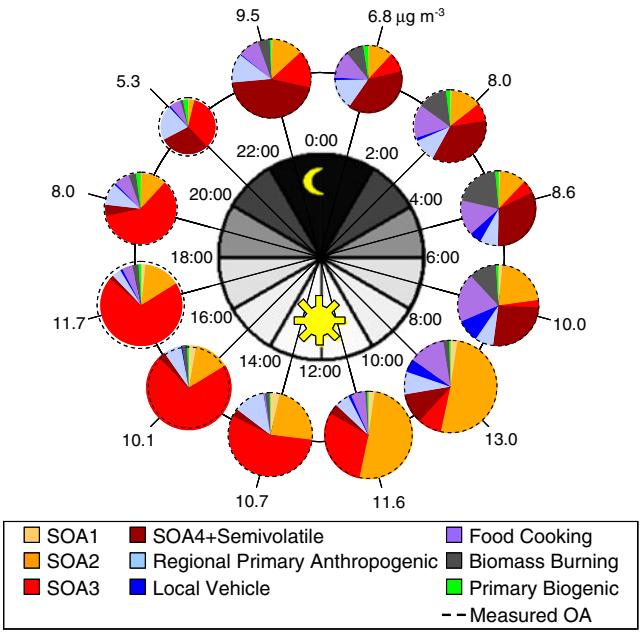

Fig. 3 Average diurnal concentrations of TAG-derived PMF factors over the summer focus period during the summer 2005 study in Riverside, California (inland Los Angeles area). *PM_1 organic aerosol mass concentrations* are labeled outside the pie chart ring, and *time of day* is labeled inside the pie chart ring. (Fig. 15 in Ref. [69], reprinted with permission)

sources due to transported and local anthropogenic pollution, transported and local biogenic emissions, and a local marine or dairy source, for a summer 2004 dataset in coastal Nova Scotia, Canada [70]. However, PCA apportions variance, and the resulting time series and mass spectra of the factors can contain negative values. PCA is therefore fundamentally different from PMF, which apportions the mass directly into nonnegative solutions that are physically meaningful.

Dreyfus et al. [34] used PMF to study the time trends of 60 organic molecular and fragment ions measured with a photoionization aerosol mass spectrometer (PIAMS) with a time resolution of a few minutes. Six factors were identified and linked to POA sources, including diesel exhaust, car emissions/road dust, and meat cooking. The mass contributions of individual sources were subsequently estimated, similarly to Williams et al. [69], by combining the PMF results and the EC/OC data. Figure 4 shows the results of a factor attributed to meat cooking aerosol, based on:

1. the mass spectrum that shows prominent peaks at *m*/*z* values corresponding to the molecular ions of palmitic, linoleic, stearic, and oleic acids, all of which are tracer compounds for meat cooking (Fig. 4a);
2. the diurnal pattern that shows two characteristic peaks consistent with typical mealtimes (Fig. 4b); and
3. the wind rose plot which shows features consistent with the locations of cooking facilities near the site (Fig. 4c).

Similar diurnal dependences were observed for cooking aerosol factors identified by factor analysis of AMS datasets at several locations, as discussed above for New York City [39, 51, 66, 71, 72].

Evaluation and selection of PMF solutions

Although a major objective of multivariate factor analysis is to explore underlying covariation of variables in a dataset to extract physically meaningful factors that can be related to distinct sources, processes, and/or properties, the solution algorithms provide only mathematical solutions that require careful evaluation and interpretation. In this section the various steps in PMF analysis are illustrated within the framework of AMS data. Ulbrich et al. [38] conducted a thorough assessment of PMF modeling of AMS data and discussed in detail several technical aspects of the analysis, including error matrix preparation, data pretreatment, selections of the optimum number of factors (*P*) and rotational forcing parameter (FPEAK), and evaluation of PMF solutions. These steps are summarized in Table 1. These authors also reported the development of an Igor-based (WaveMetrics, Lake Oswego, OR, USA) open-

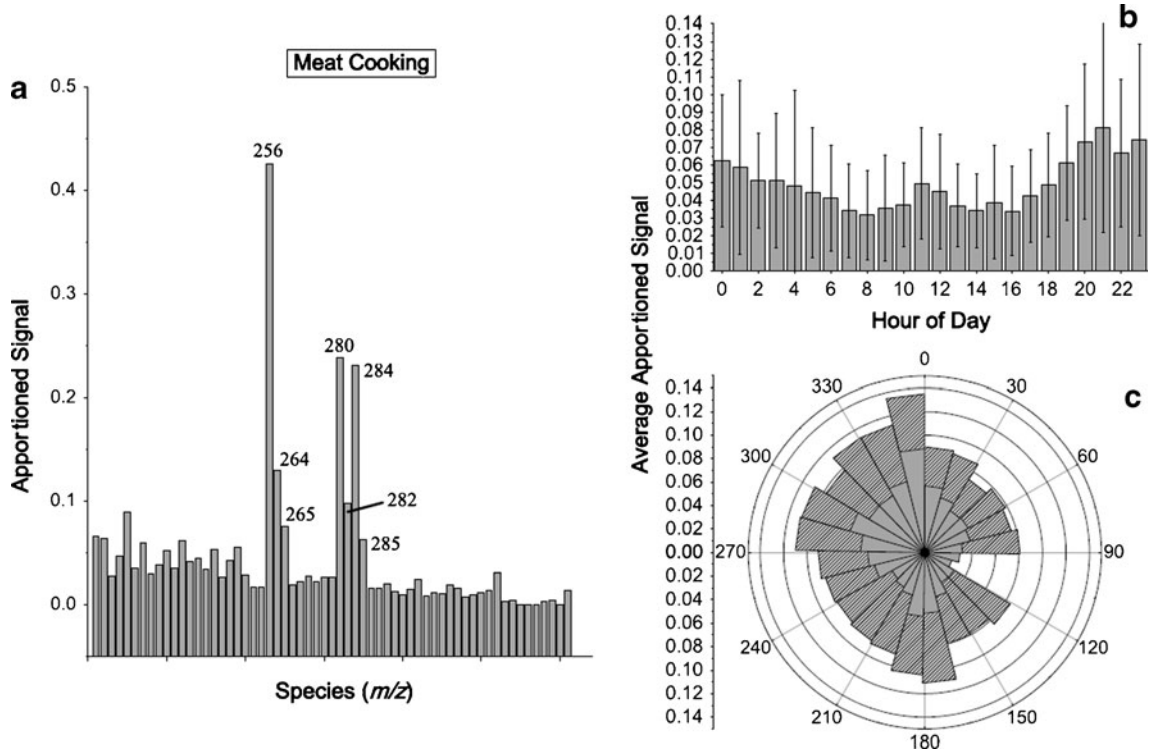

Fig. 4 A meat cooking factor obtained via PMF analysis of a photoionization aerosol mass spectrometer (PIAMS) dataset acquired in fall, 2007, in Wilmington, Delaware, USA. (**a**) Apportioned signal vs. *m*/*z*; (**b**) diurnal profile (*error bars* show the standard deviation of the range of values at each time point); (**c**) wind rose plot for high-impact periods (*outer plot*) and all data points (*inner plot*). (Fig. 3 in Ref. [34] Copyright 2009. Elsevier, reprinted with permission)

source PMF Evaluation Tool (PET, available at http://cires.colorado.edu/jimenez-group/wiki/index.php/PMF-AMS_Analysis_Guide#PMF_Evaluation_Tool_Software) that enables systematic probing of the PMF solution space, automated batch analyses, and user-friendly visualization and intercomparison of the solutions and residuals [38].

Figure 5 shows a summary of key diagnostic plots useful for evaluating the final PMF results from an AMS dataset, which should also be useful for other aerosol mass spectrometers. An important first step of the PMF analysis is to decide the optimum number of factors that "best" explain the data. Ulbrich et al. [38] demonstrated that the trend of the PMF quality-of-fit parameter (Q) changing with regard to the number of factors can be useful to identify the minimum number of factors (Fig. 5a). A large decrease in Q/Q_{exp} with the addition of a factor indicates that the additional factor is able to explain a significant fraction of the variation in the data unaccounted for by the others. In addition, examining the Q/Q_{exp} contributions per column (m/z) or per row (t) in the matrix may help identify individual m/z values or time steps that affect the model lack-of-fit most strongly (Fig. 5h, i). Depending on the causes, and especially when non-physical effects (e.g., instrumental issues) cause extraneous variability that interferes with the PMF identification of the real components, the corresponding variables may be properly downweighted or even removed to reduce disproportionate effects on the fitting outcome [74].

The model residual time series (i.e., the difference between the summed measured mass spectrum and its modeled approximation; Fig. 5g) is particularly useful for evaluating the solutions of PMF and related methods [38, 41]. The presence of time-dependent structure in the residual suggests the need for additional factor(s) for better fitting. However, for ambient datasets (Fig. 5g), it is common that substantial structure remains in the residual time series after all physically meaningful factors have been assigned. A main reason for this is true variations in the spectra of the factors, which cannot be captured with a reasonable number of components given the assumption of constant spectra in bilinear models including PMF [38]. As Ulbrich et al. [38] pointed out, this assumption of constant factor mass spectra in bilinear methods can limit the retrievability of small factors from AMS datasets.

When P is chosen, the stability, uniqueness, and interpretability of the factor solutions should be checked. The rotational ambiguity of the solutions may be explored by changing the FPEAK [48]. Both Lanz et al. [51] and Ulbrich et al. [38] discussed the variations in factors vs.

Table 1 Steps for preparing and choosing the best solution from PMF analysis of AMS datasets

	Data matrix	Uncertainty matrix	Ref.
1. Calculate data and error matrices	X	X	[73]
2. Further data and error treatment			
2a. Apply minimum error		X	[38]
2b. Remove anomalous spikes, if desired	X	X	[41]
2c. Smooth data, if desired	X	X	[38]
2d. Downweight low-SNR data		X	[38, 74]
2e. Downweight repeated information (m/z 44 and related m/z values)		X	[38]
3. Run PMF for a range of number of factors (P) and random starts (SEEDs). Examine Q/Q_{exp} vs. P in solution (Fig. 5a). A steep change in slope indicates the minimum P to consider for a good solution			[38]
3a. Examine solutions from different random starts for each P. Sort solutions by Q/Q_{exp} values and compare the factors in each solution			[52, 62]
Are there multiple types of solutions (representing local minima in the solution space)?			
If not, proceed and use seed 0			
If yes, can any of the solution types be excluded because the factors are not physically meaningful? Proceed, exploring seeds that have solutions with physically meaningful factors			
3b. Try to determine the optimum number of factors by examining multiple criteria:			[38]
Look for correlations between factor time series and time series of external tracers			[39]
Look for correlations between factor time series and time series of individual m/z values or ions			
Consider factor diurnal profiles, meteorological data, etc.			
Examine factor mass spectra for tracer ions and fragmentation patterns			
Look for signs of "split" factors, considering the correlation of mass spectra and time series of factors in the same solution. After identifying factors that may have split, explore solutions with more factors to check for new, physically meaningful factors			
3c. Examine solution Q contributions and residuals			[38]
Do the residuals and Q values summed to form time series or mass spectra show periods or m/z values that do not fit well? Is this because the solution needs more factors, because the data do not fit the model of constant spectra for a given component, or because of instrumental drift, etc.?			
Are the distributions of the scaled residuals (x_{ij}/σ_{ij}) for each m/z approximately Gaussian, centered around 0, with a reasonable standard deviation?			
4. For the best solution chosen from step 3, run PMF for a range of FPEAKs such that the range of Q/Q_{exp} values is at least 3% above the minimum Q/Q_{exp}			[38, 51]
4a. Exclude from further consideration solutions that have unrealistic mass spectra and/or time series			[38]
4b. Does changing FPEAK change the solution in a way that would change the interpretation of the factors from step 3, or do these solutions just represent rotational ambiguity in the solution?			[38]
If the interpretation changes, choose the most representative solution and support this choice			
If the differences represent rotational ambiguity, choose the solution at FPEAK = 0			
5. Conduct bootstrapping analysis on the final solution from step 4 to estimate uncertainty in the solutions			[38]
6. Make and examine key diagnostic plots			[38]
6a. Q/Q_{exp} vs. varying P (e.g., Fig. 5a)			
6b. Q/Q_{exp} vs. FPEAK for the best P (e.g., Fig. 5b)			
6c. Fractions of OA factors vs. FPEAK for the best P (e.g., Fig. 5c)			
6d. Correlations among PMF factors for the best P (e.g., Fig. 5d)			
6e. The box and whiskers plots of scaled residuals as a function of m/z for the best P (e.g., Fig. 5e)			
6f. The time series of the measured OA concentration and the reconstructed organic mass (= sum of all factors for the best P; e.g., Fig. 5f)			
6g. The variations of the residual (= measured − reconstructed) of the fit as a function of time (e.g., Fig. 5g)			
6h. The time series and mass spectra of total residuals and Q contribution for the best P solution (e.g., Fig. 5h, i)			
6i. Comparisons of the $P - 1$, P, and $P + 1$ solutions for the acceptable FPEAK, where P is the best solution			

FPEAK (Fig. 5b). The robustness of the solution may also be examined by running the PMF algorithm from different random starting points (SEED parameter). Variations in different plausible solutions corresponding to different

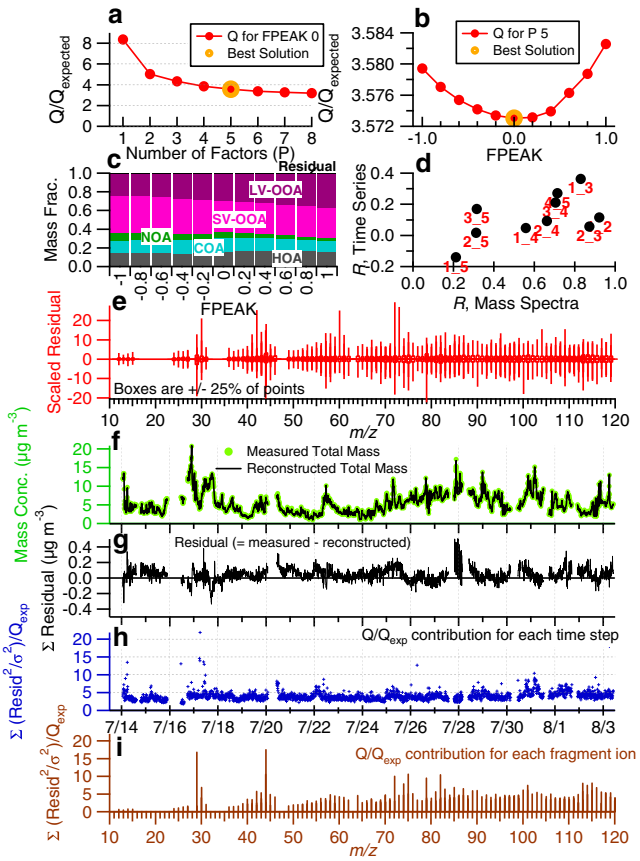

Fig. 5 Summary of key diagnostic plots of the PMF results for an HR-ToF-AMS dataset acquired in New York in 2009 (the PMF results are given in Fig. 2): (**a**) Q/Q_{exp} as a function of number of factors (P) selected for PMF modeling. For the five-factor solution (i.e., the best P): (**b**) Q/Q_{exp} as a function of FPEAK, (**c**) fractions of OA factors vs. FPEAK, (**d**) correlations among PMF factors, (**e**) the box and whiskers plot showing the distributions of scaled residuals for each m/z, (**f**) time series of the measured organic mass and the reconstructed organic mass (= LV-OOA + SV-OOA + NOA + COA + HOA), (**g**) variations of the residual (= measured − reconstructed) of the fit, (**h**) the Q/Q_{exp} for each point in time, and (**i**) the Q/Q_{exp} values for each m/z. (Adapted from Ref. [39])

FPEAK or SEED values may be evaluated to determine the uncertainties of the PMF solution [38, 49, 75]. In addition, the uncertainty of the solution corresponding to a given P and FPEAK can be analyzed quantitatively using bootstrapping analysis [38]. Finally, the interpretability of the OA factors should be evaluated on the basis of their mass spectral features and temporal variation patterns (details are given in the section "Interpretation of the extracted OA factors").

Interpretation of the extracted OA factors

The objective of interpreting the solutions of PMF and similar methods is to identify and validate the relationships between OA factors and distinct emission sources, physicochemical properties, and atmospheric processes. The interpretability of the OA factors is also an important criterion for evaluating the quality of the multivariate analysis. The interpretations of the OA factors are usually based on the following considerations:

1. the temporal correlations of factors with tracer species representative of specific emissions and processes;
2. the mass spectral features of each factor, for example peak distribution patterns, signature fragments, and oxidation state;
3. the repetitive temporal or diurnal variation patterns that are indicative of specific human activities or meteorological patterns (for example traffic rush hours, dilution because of the increase of the planetary boundary layer, cooking emissions during mealtimes, photochemical production of secondary species, etc.);
4. the estimated size distributions of OA factors (or tracer ions) and their evolution patterns;
5. information regarding airmass trajectories and locations of upwind source regions; and
6. other collocated observations that enable the isolation of special cases (e.g., new particle formation and growth events identified according to scanning mobility particle sizer measurements [40] and well-defined SOA growth events [42]).

The correlations between the time series of OA factors and those of independent external tracer species (i.e., species not included in **ORG**) are especially important for addressing the physical meaning of the OA factors. High time resolution of the OA factors greatly facilitates interpretation of their physical meaning, and this is probably the most important reason for the rapid acceptance of factor analysis by the research community. The fast measurements capture the dynamic variations caused by true changes in aerosol sources and transport, while minimizing the uncertainties caused by apparent correlations with tracer species because of longer time averages in, e.g., filter analyses.

An important step of factor interpretation is to compare the extracted factor spectra with reference spectra sampled from various source types, a large number of which have been published in the literature [14, 76–84] and are publicly available in the AMS Spectral Database at http://cires.colorado.edu/jimenez-group/AMSsd/. An especially useful set of spectra for ambient HOA, SV-OOA, LV-OOA, and OOA were reported by Ng et al. [14] by averaging OA factors determined from a large number of ambient AMS datasets. The similarity between two mass spectra (or two time series) can be evaluated using Pearson's R, the coefficient of determination (R^2), or the uncentered correlation coefficient. In addition to comparisons of the full mass spectra, examining the correlations among peaks

above *m/z* 44 can avoid biases caused by small *m/z* ions that generally dominate the mass spectra [38, 51]. At times, however, unrealistic factors can have spectra which look similar to those in the database, so this criterion is not sufficient for supporting the identification of a factor [38]. The presence of key marker ions, for example *m/z* 44 for OOA or *m/z* 60 for BBOA, is another useful criterion for factor interpretation [14].

Examples of established tracer-factor relationships are discussed here. A large number of studies have demonstrated the dominant association of OOA with SOA, based on observations that OOA generally correlates well with:

1. secondary inorganic aerosol species—sulfate, nitrate, and/or non-refractory chloride (e.g., ammonium chloride) [6, 38–40, 42, 66, 72, 85, 86];
2. water-soluble organic carbon (WSOC) concentrations [87]; and
3. gas-phase photochemically-produced species such as odd oxygen ($O_x = NO_2 + O_3$) [88, 89] or glyoxal [42].

The association of HOA with OA from primary sources has been supported by collocated measurements of tracer species associated with combustion emissions, including CO, NO_x, polycyclic aromatic hydrocarbons (PAH), and black or elemental carbon (BC/EC). Specific evidence includes:

1. good correlation between the concentrations of HOA and combustion tracer species [6, 38–41, 64, 66, 72, 85, 128];
2. estimated emission ratios of HOA against EC, NO_x, and CO (i.e., HOA/EC, HOA/NO_x, or HOA/CO) consistent with source measurements [6, 39, 40, 43, 64]; and
3. consistency between HOA and POA concentrations estimated using tracer-based approaches, for example the EC tracer method, the CO tracer method, and the chemical mass balance (CMB) model using organic molecular markers [40, 43, 63, 64].

In addition, the mass spectra of HOA from various studies generally show fragmentation patterns characteristic of long-chain hydrocarbons and are very similar to those of diesel exhaust, lubricating oil, and freshly emitted traffic aerosols observed in urban areas [14, 41]. HOA appears chemically reduced with average oxygen-to-carbon (O/C) ratio typically less than 0.1 [13, 14, 31, 39, 40].

The identification of OA factors associated with biomass burning (i.e. BBOA) has been supported by correlations of this factor with biomass burning emission tracers (e.g., acetonitrile, levoglucosan, potassium, and non-fossil EC), elevated peaks at *m/z* 60 ($C_2H_4O_2^+$) and 73 ($C_3H_5O_2^+$) in the mass spectra of the factor, and model dispersion analyses from the locations of known forest fires [22, 65, 79, 80, 83, 85, 90].

COA has been identified in several studies as discussed above. In other studies, especially those using UMR data, it is often not separately identifiable and may be part of the HOA and/or BBOA factors [72], because of the similarity b the COA spectra (especially the UMR spectra) and the HOA and BBOA spectra [83]. External tracers of food cooking are not usually available, but the diurnal profiles of the factors show that HOA peaks during rush-hour periods whereas COA peaks during typical meal times. Sun et al. [39] suggested that the $C_5H_8O^+$, $C_6H_{10}O^+$, and $C_7H_{12}O^+$ ions in the HR-AMS spectra may potentially be useful as AMS-spectral markers for COA. In addition, the COA and HOA factors could be differentiated on the basis of the signal ratio of *m/z* 55 to *m/z* 57 as the COA spectrum tends to show substantially higher *m/z* 55 to 57 ratio [39, 83, 84].

Nitrogen-enriched OA or local OA factors (NOA or LOA) have been reported in several studies [39, 58, 64]. The NOA or LOA mass spectrum has important contributions from many nitrogen-containing fragments not observed in other factor mass spectra, and consistent with reduced nitrogen species such as amines, amides, or nitriles [39]. Although no external tracers have been identified to help link NOA components to a particular source or process, NOA factors tend to have spiky time series and are, therefore, likely be the results of more local emissions, i.e., if NOA was emitted or produced farther away, it would disperse in the atmosphere and have a smoother time series.

Analysis of the air mass trajectory histories and comparisons with the results of other source apportionment techniques may provide further support for the interpretation of the OA factors [22, 41, 63, 72, 91, 92]. In addition, the chemically-resolved size-distribution data from the AMS are valuable for elucidating the sources and processes of OA factors. For instance, the frequently-observed similarity between the size distribution of *m/z* 44 (AMS tracer for OOA) and sulfate supports the association of OOA with SOA [39, 40, 43, 93–96]. Zhang et al. [40] estimated the size distributions for OOA and HOA based on measured size distributions of *m/z* values 44 and 57 and the mass spectral patterns of OOA and HOA using a UMR AMS dataset acquired in Pittsburgh in 2002, as shown in Fig. 6. The size distribution of HOA shows a distinct ultrafine mode that is commonly observed for primary particles from fresh combustion emissions [12, 76, 97] most prominently during morning rush hours and at night when boundary layer height is low and atmospheric dilution of the primary emissions is weak. In contrast, OOA is concentrated in the accumulation mode peaking between 400 and 600 nm (in vacuum aerodynamic diameter; d_{va} [30]) and seems to be mostly internally mixed with sulfate, a secondary inorganic species. The

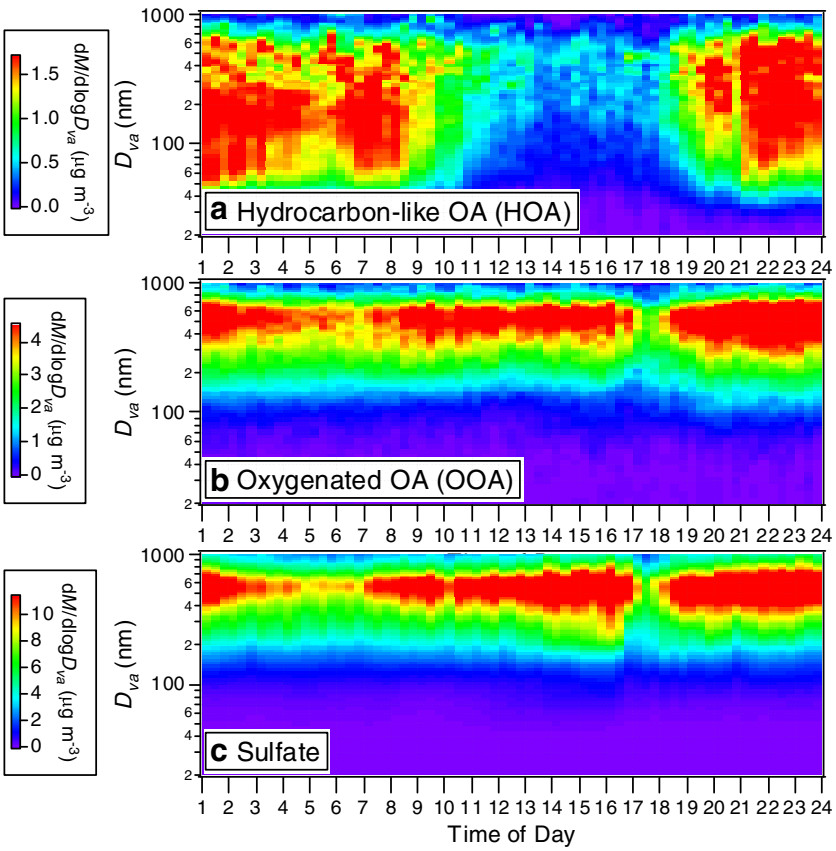

Fig. 6 Average diurnal variations of the estimated size distributions of (**a**) HOA, (**b**) OOA, and (**c**) sulfate during 7–22 September, 2002, in Pittsburgh, USA. (Figure 6 in Ref. [40], reprinted with permission)

short lifetime of ultrafine particles supports the association of HOA with more local sources. In contrast, the diurnal variations of OOA and sulfate are relatively weak, both in terms of size distributions and concentrations, indicating strong effects from regional sources and processing. The fact that OOA and sulfate correlate well both in concentration and size distribution strongly supports the dominant secondary contribution to OOA, and may also be indicative of the effect of cloud processing on SOA production similar to that on sulfate. Furthermore, an additional piece of evidence for the secondary nature of OOA is that the evolution pattern of OOA size distribution during an intense new particle formation and growth event clearly indicates growth of OOA via surface condensation (i.e., gas to particle conversion) [40, 95].

Advanced factor analysis of aerosol mass spectra

In addition to standard PMF bilinear modeling, more advanced factorization approaches have been applied to aerosol mass spectrometry data to improve the specificity and interpretability of the solutions. Slowik et al. [50] successfully performed PMF analysis on a combined matrix of the OA mass spectra acquired with a UMR AMS and the mass spectra of volatile organic compounds acquired with a proton-transfer-reaction mass spectrometer (PTR-MS) during winter in Toronto. The uncertainties used in PMF for each instrument were scaled to result in similar weights in the PMF analysis. Six factors characteristic of charbroiling, traffic, aged SOA, local SOA, oxygenated POA, and a local point source were identified, with information on the temporal and source profiles of both OA and VOCs for each factor (Fig. 7). According to the authors, PMF analyses of separated AMS or PTR-MS datasets were not able to identify as many factors, because of effects (e.g., collocated emissions and meteorological variations) that blur the distinctions between primary and secondary species in the same phase, thus enhanced variance in the unified AMS/PTR-MS dataset was thought to have enabled the distinction of more similar factors. Another important advantage of incorporating the VOC data was the simultaneous and coherent apportionment of VOCs to the same emission sources and atmospheric processes represented by the OA factors. The availability of both VOCs and OA profiles also facilitated the interpretation of the factors. However when atmospheric aging is more important than in this winter Northern latitudes dataset, the VOC profiles are strongly distorted by the photochemistry in a time-dependent manner which is inconsistent with PMF's assumptions [98], and a similar joint AMS + PTR-MS analysis has not been reported under those conditions.

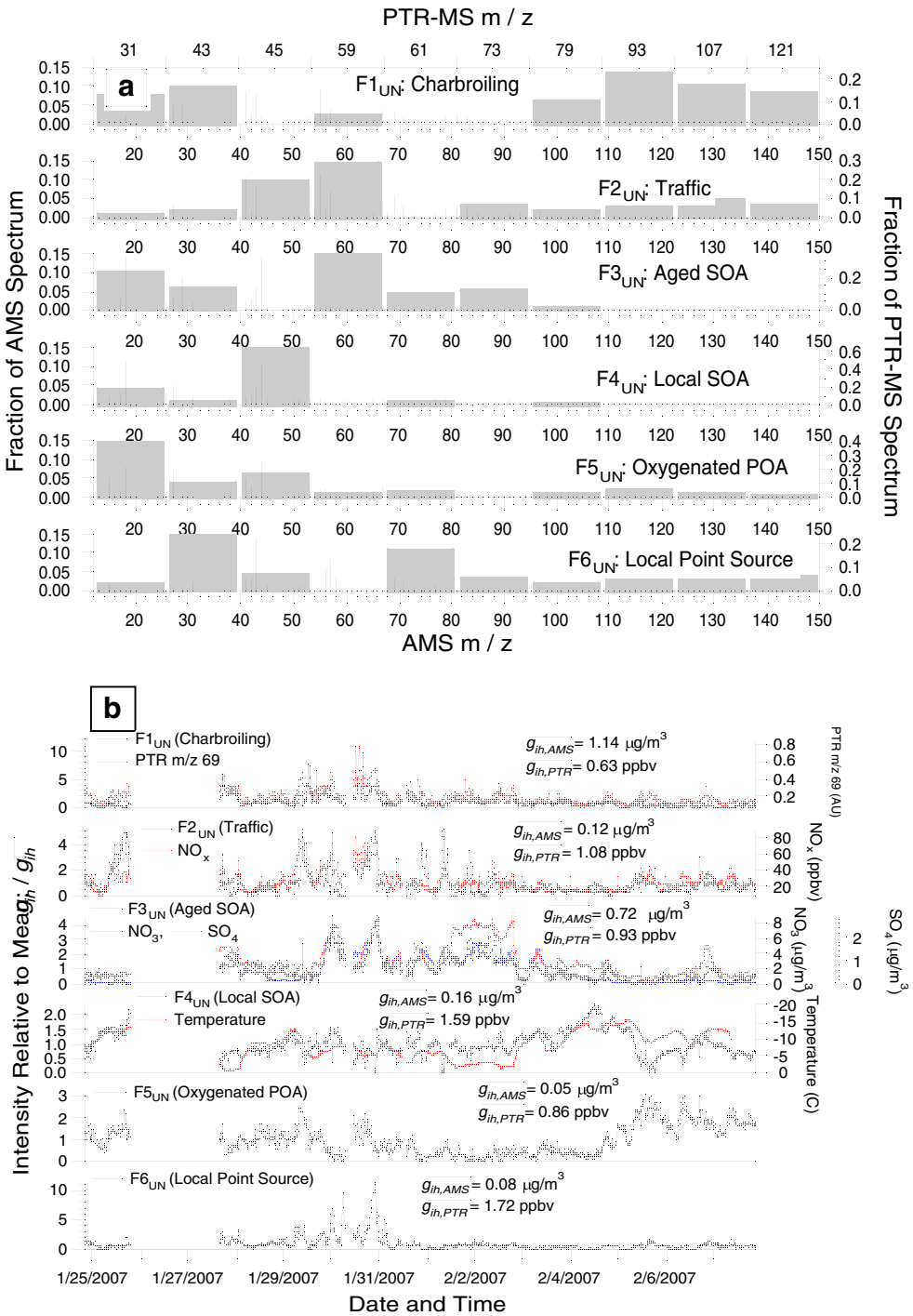

Fig. 7 PMF analysis results of a unified dataset of AMS and PTR-MS measurements acquired in winter, 2007, from Toronto, Canada. Mass spectra (**a**) and time series (**b**) of the PMF factors (*black traces, left axis*) and selected tracer species (*colored traces, right axis*). The time series of PTR-MS *m/z* 69 is plotted in arbitrary units. (Figure 11 in Ref. [50], reprinted with permission)

Lanz et al. [85] factored an AMS dataset collected in Zurich during a period dominated by wintertime inversions. Because of the high residence time of air masses, species from different sources co-varied to such a extent that standard PMF analysis could not separate physically meaningful source profiles. This is because of an inherent limitation of the PMF and similar algorithms, which have difficulty resolving factors that are either too similar in mass spectra or time series. To enable the extraction of physically meaningful factors, Lanz et al. [85] incorporated estimated source profiles and solved the bilinear model by use of multilinear engine (ME-2) software [53, 99]. ME-2 uses a different algorithm that can solve the "standard" bilinear PMF, and many other factorization models. The introduction of a-priori source profiles in the fitting, which is not a requirement when using ME-2, can be viewed as a hybrid of the PMF and the chemical mass balance (CMB) model, which determines the relative source contributions

of OA factors using fixed known source and/or mass spectral profiles [100]. Using the MS of POA from diesel bus emission experiments as the first-guess a-priori profile and allowing the profile to partially deviate from the a-priori one, three distinct OA factors, including an OOA mostly representing SOA, a factor representing BBOA from wood combustion, and a traffic-related HOA, were identified from the Zurich dataset [85]. Figure 8 shows how the spectral similarities between hybrid and reference OA factors change as a function of the regularization parameter ("*a*-value"), which defines the degree of constraint of the HOA factor. An "*a*-value" of 0 means that the a-priori profile (i.e., the mass spectrum of the HOA factor) is not allowed to change during iterative fitting. This figure indicates that a=0.6 is a good compromise for this dataset because it allows flexibility in the solution while the factor spectra still show high similarity to reference profiles.

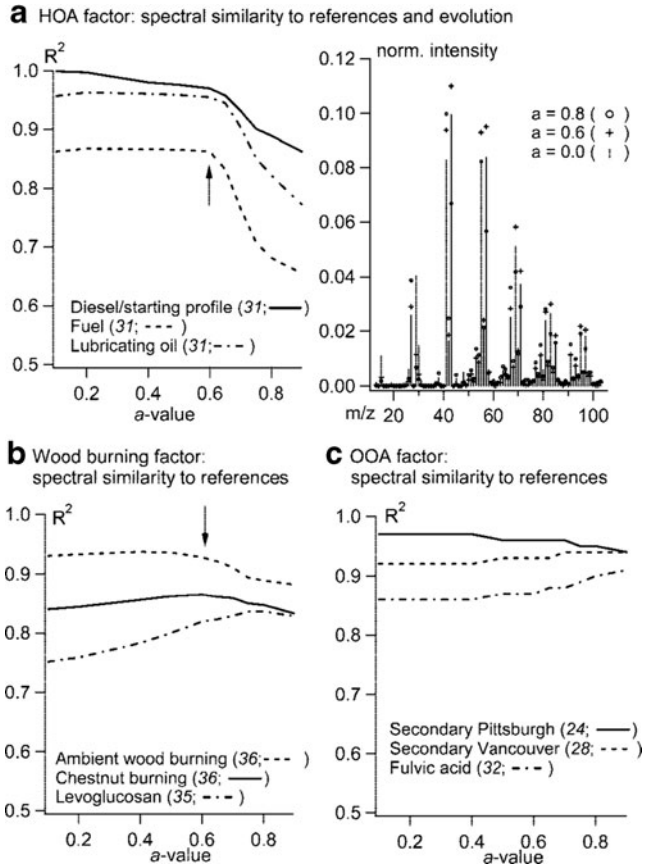

Fig. 8 Results from hybrid PMF-CMB analysis of an AMS dataset acquired during wintertime in Zurich, Switzerland. (**a**) Spectral similarity (R^2) of the retrieved HOA factor spectrum to the reference mass spectra representing urban combustion POA and the changes of the HOA factor as a function of the regularization parameter ("*a*-value"), which is the degree of deviation from the initial spectrum allowed for the HOA factor. Similarity (R^2) of the mass spectra of wood burning (**b**) and OOA (**c**) factors to reference mass spectra representing BBOA and OOA, respectively, as a function of the "*a*-value". The *values in the parentheses before the line symbols* are the numbers of references cited in that paper. (Fig. 3 in Ref. [85]. Copyright 2008. American Chemical Society, reprinted with permission)

DeCarlo et al. [62] reported the application of PMF to flight data from the Mexico City region, when fresh urban and biomass burning emissions mixed and underwent very strong photochemical aging. In this case PMF identified HOA, BBOA, SV-OOA, and LV-OOA factors, but evidence such as ratios of components to tracer species indicated the combined effect of urban and biomass burning sources in several factors. These authors applied a "postprocessing" step to the PMF output in order to apportion fractions of some of the components to urban and biomass burning sources, based on ratios of components to tracers and the observed variations between days with and without the effect of intense biomass burning.

Ng et al. [101] first reported the application of full CMB modeling to AMS data using experimentally determined aerosol source profiles as fixed input mass spectra. A main advantage of CMB is the ease of solution using a simple linear decomposition algorithm, which can be performed on data acquired in real-time without waiting for a sufficiently large dataset for PMF analysis. Consistent with the discussion of the previous study, CMB may also be able to separate factors that correlate strongly in time, because of the collocations of emission sources and/or meteorological effects. However, CMB is only suitable when appropriate and complete source profiles are available. Average mass spectral profiles for ambient HOA, OOA, LV-OOA, and SV-OOA have recently been determined by Ng et al. [14] by averaging those observed in many studies. When performing CMB of several datasets, the component concentrations were typically within 30% of those determined using full PMF analysis.

Recently, Ulbrich et al. (2011) reported for the first time the application of two three-dimensional factorization models ("PARAFAC" and "Tucker-1"), solved with the PMF3 [102] and the ME-2 [53] algorithms, to the size-resolved mass spectral dataset from an HR-ToF-AMS. In contrast with quadrupole AMSs, AMS systems using time-of-flight mass spectrometers are able to acquire full mass spectra of OA for each particle size, i.e., to determine the size distributions of each individual m/z in the spectrum. By applying 3D factorization, Ulbrich et al. [52] determined the size distributions of four OA factors HOA, OOA, BBOA, and a nitrogen-enriched factor (NOA or LOA). These factors were consistent with source measurements and previous estimates of the component size distributions, and enabled identification of cases when, e.g., HOA-containing particles grew strongly because of the

condensation of secondary species. An important aspect of this work is that the 3D factorization approaches, methodologies, and considerations are generally applicable to multi-dimensional mass spectral datasets such as those from thermal-desorption mass spectrometers or hyphenated aerosol mass spectrometers (e.g., coupled with chromatography for chemical separation).

Insights into organic aerosol sources, processes, and properties

Overview of OA factors and their spatial distribution

An important advantage of factor analysis is that it reduces the extremely complex OA composition to a limited number of chemically and physically meaningful components that can be linearly combined to reproduce the observed time and chemical variations in ambient OA. A broad overview on the chemistry, variability, and evolution characteristics of atmospheric submicron OA is emerging from the analysis of a large number of AMS datasets acquired from worldwide locations. Zhang et al. [6] performed an integrated analysis of 37 AMS datasets collected from three continents (North America, Europe, and Asia) and concluded that oxygenated OA species (surrogate for SOA) are ubiquitous and dominant in the Northern Hemisphere. PMF analyses of these worldwide AMS datasets further led to the compilation of a global picture on the oxidation states and dynamic evolution of multiple OA factors [13, 14, 103], which complements those obtained by use of other techniques [104]. In addition, two recent studies discussed the spatial distribution of aerosol chemical composition and the evolution of OA in Europe based on measurements of aerosol chemical composition across the continent [18, 56].

Figure 9 shows a summary of the results from PMF analysis of 43 worldwide datasets. Usually 2–5 OA components were extracted for each study, consistent with those discussed above. It is important to note that the mass spectrum corresponding to any given factor is similar but not identical across multiple sites [14]. From Fig. 9 it is clear that the average mass concentrations and compositions of submicron particles vary substantially across sites. Whereas both HOA and OOA loadings decrease with distance from urban sites, the decrease of HOA is much more pronounced and the average OOA concentrations are often of the same order in various atmospheric environments [6]. The relative contribution of OOA evolves from an average of 63% of the organic matter in submicron particles at urban locations to usually more than 90% in rural and remote locations [6]. This difference in relative contributions reflects the spatial differences in the sources of HOA and OOA, respectively. Primary sources of OA (particularly HOA) are highest in urban centers and are quickly diluted when advected away from their sources. SOA is thought to be produced on time scales of a day or more in amounts that greatly exceed the urban HOA [10], so OOA corresponds not only to locally produced SOA but also SOA produced downwind of polluted areas and thus has a larger footprint (i.e., over wider regions). In addition, the source region for SOA precursors has additional sources with a more regional footprint, including regional biogenic emissions and biomass burning emissions.

Figure 10 provides a clear example of this difference in HOA and OOA source footprints. The data points shown in the figure correspond to AMS and CO_2 (as a combustion tracer) gradient measurements downwind of a highway source (I-93 in Somerville, MA, USA) reported in Canagaratna et al. [12]. The measurements were made aboard a mobile laboratory equipped both with instruments that responded rapidly to aerosols and with gas phase instruments, with AMS measurements performed every few seconds [105]. The OOA concentrations were not significantly affected by the highway source and no gradient was observed. The gradient in HOA concentrations, on the other hand, mirrors the decreasing CO_2 concentrations, and is indicative of the effect of dilution as the pollution plume is mixed with ambient air.

When HOA particles mix in with background air, they can also become coated with background secondary aerosol material (i.e. species such as ammonium sulfate, ammonium nitrate, and secondary organic species). Cross et al. [106] used PMF analysis of single particle mass spectra measured with the single particle light-scattering AMS (SP LS-AMS) to show that HOA particles, which were likely to be emitted as POA from a near-by highway, quickly become internally mixed with secondary inorganic and organic compounds after emission. Canagaratna et al. [12] determined the effective densities of the single particles examined in the Cross et al. [106] study and confirmed that they were consistent with their PMF-based classification as “pure” and “mixed” HOA single particles. Taken together, these results indicate that coating processes are likely to modify HOA particles and lead to an evolution of their properties and effects with time.

OOA subtypes and interpretation

In environments in which chemical variation in the OOA is not significant, either because of lack of change in the meteorological conditions (e.g., ambient temperature), source effect, or photochemical age distribution, only one OOA is extracted [14]. However, in environments where

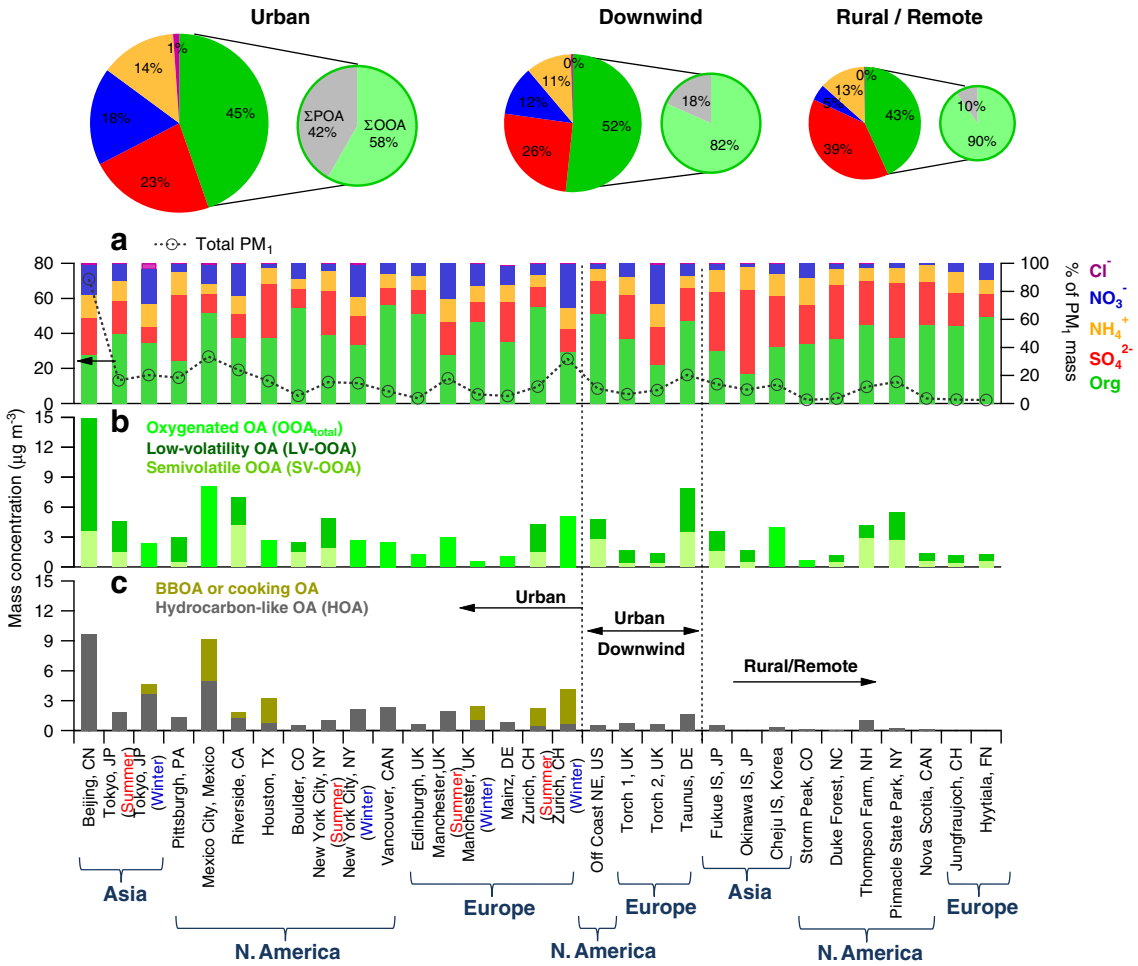

Fig. 9 (**a**) Average total mass concentration and composition of PM_1; (**b**) average mass concentrations of oxygenated OA factors (including OOA_{total}, LV-OOA, and SV-OOA); and (**c**) average mass concentrations of primary OA factors (including HOA, BBOA, and COA) in PM_1 at individual sites in northern hemisphere (Plotted using data from Zhang et al. [6], Jimenez et al. [13], and Ng et al. [14])

OOA is subject to continuous evolution, several OOA factors which represent the end points of a relatively continuous chemical variation arising from different levels of aging are observed [14, 18, 39, 91]. In many cases LV-OOA and SV-OOA, which appear to be surrogates for regional/more-aged and fresher/less-aged SOA, respectively, are observed [38, 39, 51, 62–64]. The volatilities of OOA subtypes were confirmed by measurements of aerosol volatility using a thermodenuder to remove semivolatile species in a temperature-dependent manner [58, 107] or inferred according to their correlations with ammonium sulfate (suggesting low volatility) and with ammonium nitrate and chloride (suggesting semi-volatility) [101]. Whereas the separation of LV-OOA and SV-OOA is frequently reported, especially during summer time, a few studies have reported OOA factors with different O/C ratios but similar volatility [66, 72, 91]. In these cases the SV-OOA and LV-OOA terminology is not appropriate, and that of more and less oxygenated OOA (MO-OOA and LO-OOA) should be used instead. The different OOA factors may also reflect other sources or processing factors that cause differences in the spectra of groups of oxygenated species. In particular, at some locations affected by distinguishable

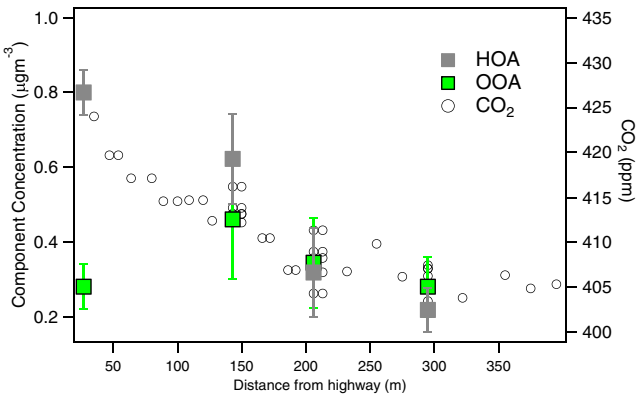

Fig. 10 Profiles of the concentrations of HOA, OOA, and CO_2 (as an indicator of dilution) as a function of distance downwind of a major highway (Interstate 93) in Somerville, MA, USA. (Adapted from Fig. 4a in Ref. [12])

biogenic and anthropogenic impact periods, for example Chebogue Point, Nova Scotia [82], Whistler Peak, British Columbia [108], and the Egbert rural site 70 km north of Toronto [67], the less oxidized OOA component seems to be associated with biogenic emissions (i.e., a surrogate for biogenic SOA) whereas the more oxygenated OOA seems to be associated with air masses transported from polluted regions (i.e., a surrogate for anthropogenic and anthropogenically controlled biogenic SOA).

Ng et al. [14] studied the aging of OA components in the atmosphere by examining differences between the mass spectra of LV-OOA and SV-OOA on the basis of two main OOA ion fragments at m/z 44 (CO_2^+) and m/z 43 (mostly $C_2H_3O^+$). The composition differences between the OOA components are reflected in the different intensities of these two ions. As shown in Fig. 11, the LV-OOA component spectra have a higher f_{44} (ratio of m/z 44 to total signal in the component mass spectrum) and lower f_{43} (defined similarly) than SV-OOA. When f_{44} (a surrogate for O/C ratio and an indicator of photochemical aging) is plotted against f_{43}, ambient OOA components lie within a well-defined triangular region. The different OOA components in Fig. 11 offer snapshots of the continuum of evolving OOA properties in ambient aerosol. Morgan et al. [18] have shown a similar continuum of AMS OOA components in aircraft measurements over Europe. The less aged SV-OOA occupies the broader base of the triangle, which is likely to reflect the variable composition of fresher SOA formed from site-specific precursors and sources. The LV-OOA, on the other hand, occupies the narrowing top region of the triangle. LV-OOA factors are spectrally similar to fulvic acid and HULIS sample spectra [14]. This, together with the fact that high values of f_{44} in AMS spectra are indicative of acid groups [109, 110], suggests that the highly oxidized LV-OOA is likely to contain polyacidic or acid-derived moieties.

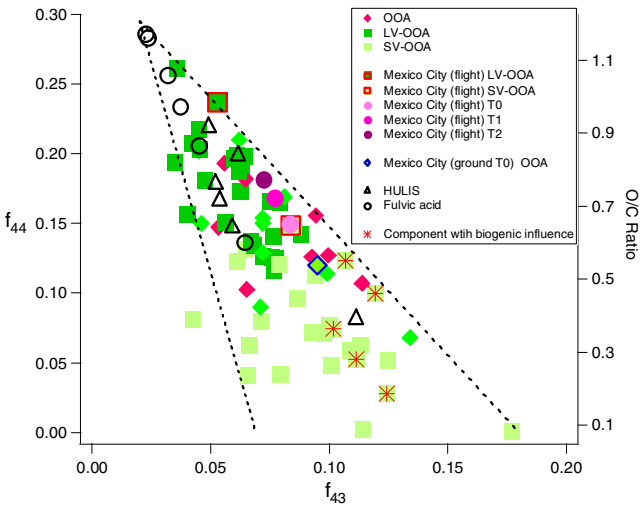

Fig. 11 f_{44} vs. f_{43} in the mass spectra of ambient OOA and various HULIS and fulvic acid samples. The *dotted lines* define the triangular space where ambient OOA components fall. The O/C ratios were estimated using the formula given in Aiken et al. [31]: $O/C = 0.0382 \times f_{44} + 0.0794$ (Adapted from Fig. 4 in Ref. [14], reprinted with permission)

A key implication of Fig. 11 is that photochemically aged SOA converges towards the same highly oxidized endpoint regardless of its original source. The common features of the OOA evolution shown in Fig. 11 potentially enable a simplified description of the oxidation of OA in the atmosphere. In fact, Jimenez et al. [13] recently used the PMF results from the datasets shown in Fig. 9 together with laboratory data to propose a two-dimensional modeling framework to map the evolution of OA. The two axes are volatility and O/C ratio. Because both of these axes can be experimentally obtained, this 2D model provides a useful framework for describing OA evolution that can be constrained by measurements [111]. O/C ratios of OA can be obtained, for example, directly from HR-ToF-AMS measurements [31] and OA factor volatilities can be measured with thermal denuders [57, 60]. Recent thermal denuder measurements indicate that ambient POA is more volatile than assumed in current models whereas ambient SOA becomes increasingly less volatile with photochemical aging, and is much less volatile than the SOA produced in smog chambers [58–60].

Recent studies have also investigated the climate-relevant properties of ambient OA factors such as their cloud condensation nuclei (CCN) activity [112–116] and hygroscopicity [13, 59, 117]. A consistent picture is emerging from the different measurements, in which fresh HOA seems to be non-hygroscopic and more oxygenated species take up more water as their O/C increases. In ship-based measurements, for example, Quinn et al. [114] showed that the variation in the critical diameter for CCN activation of marine, urban, and industrial ambient aerosol CCN activity at a supersaturation of 0.44% was correlated with the HOA mass fraction in particles with d_{va}<200 nm. Raatikainen et al. [59] have also shown that the hygroscopicity of ambient OA factors increases with increasing oxidation, which is consistent with laboratory and field measurements that show a positive correlation between hygroscopicity and overall O/C ratio [13, 117, 118].

Use of factor analysis results in regional and global modeling

Despite recent advances, better descriptions of the properties and evolution of both primary and secondary OA are needed in regional and global models [7, 10, 11]. For instance, the emission inventories and the evaporation upon dilution of POA are not well constrained [119]. Explicit

modeling of all important SOA chemical reactions and species are computationally prohibitive for large-scale models, and simpler modeling based on laboratory experiments often does not reproduce the observed concentration or dynamic evolution of SOA in polluted regions [42] while performing better over clean biogenic regions [67]. Thus, there is a continuing need for field data that can be used to accurately constrain and test model predictions.

In this context, the reduced complexity in OA factor analysis results is particularly useful for development and evaluation of air quality and climate models. OA factors extracted from factor analysis of highly time-resolved aerosol mass spectra correspond to lumped groups of molecules that are linked to each other by similar sources or processes. Thus, the observed mass concentrations and bulk composition (e.g. O/C) of these factors provide useful information for constraining theoretical predictions of the spatial and temporal evolution of key OA sources. For example, the factor analysis results of AMS data acquired from multiple stationary and aircraft platforms during the MILAGRO 2006 experiment in Mexico City have been used to evaluate box and regional scale models [120–123]. The concentrations of POA components (= HOA + BBOA) measured during the MILAGRO experiment have been used to constrain and evaluate the accuracy of the Mexico City emission inventories for primary sources [123] and several recent SOA modeling approaches have been tested against measured mass concentrations of OOA [120, 121, 124]. As an example, Fig. 12 shows a comparison of 3D model results using three different SOA modeling approaches and measurements obtained in MILAGRO [120]. The "REF" model reports the POA and SOA expected from primary emission inventories assuming that the primary emissions can be modeled as completely non-volatile and that SOA is dominated by aromatic precursors. The "ROB" model (based on Ref. [119]) treats primary emissions as semi-volatile species (SVOCs) that can evaporate from the particle phase after emissions and then react to form less volatile SOA. This model also includes intermediate volatility gas phase species (IVOCs), which are not typically measured or included in standard models, as precursors for measured SOA. The "GRI" model (based on Ref. [125]) has a similar structure to the "ROB" model, but with updated terms. Figure 12 indicates that the models that include SVOCs and IVOCs are generally in better agreement with measured factor concentrations. However, the fact that neither the "ROB" or "GRI" predictions fully captures the observations indicates the need for improved models.

The worldwide factor analysis results from the AMS data shown in Fig. 9 were recently used to constrain SOA sources using a global chemical transport model [126]. In this case the observed OOA concentrations were used to optimize the SOA source strengths from lumped biogenic, anthropogenic, and biomass burning sources. IMPROVE network measurements were then used to independently verify the optimized SOA predictions. The optimized

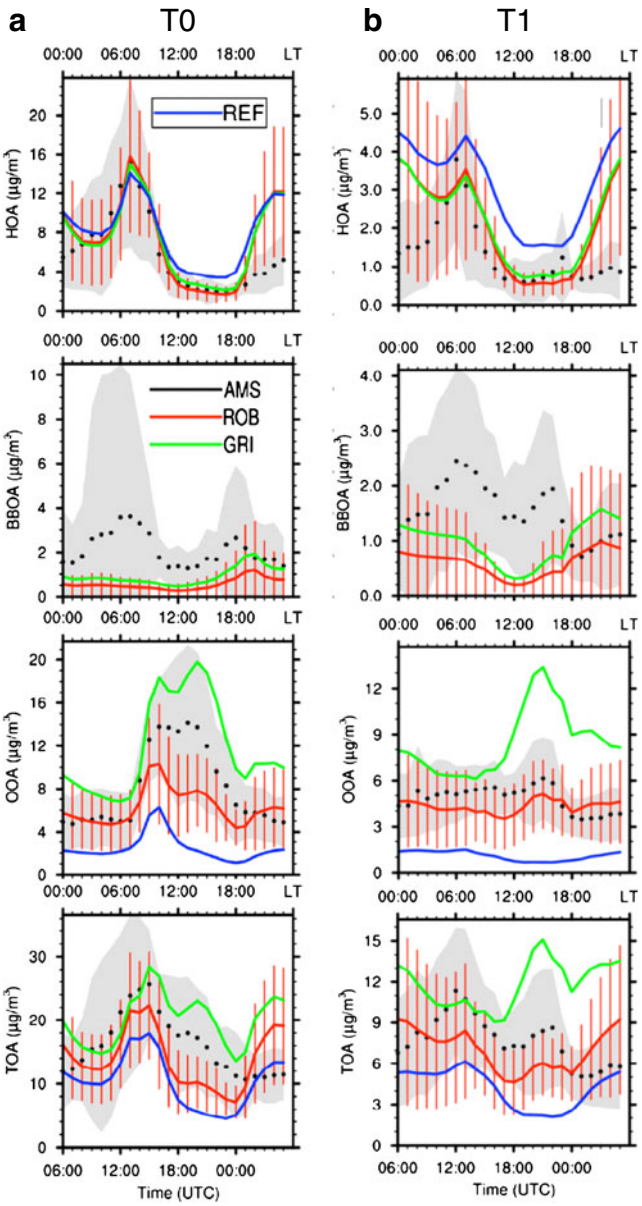

Fig. 12 Average diurnal profiles of the model simulations (*red, blue, and green lines*) and the AMS measurements (*black dots*) at the (**a**) T0 (urban center site) and (**b**) T1 (downwind suburban site) for the MILAGRO field experiment in Mexico City for the concentrations of HOA, BBOA, OOA, and total OA. The REF model case represents POA as non-volatile and only accounts for SOA formation from aromatic species. The ROB (*red line*) and GRI (*green line*) cases consider POA as semi-volatile and also account for SOA formation from semivolatile and intermediate volatility species (see text). The variability associated with average observations and ROB model predictions is given in the *gray shaded area* and *red vertical* bars, respectively. (Adapted from Ref. [120], reprinted with permission)

model yields a global SOA source of 140 Tg SOA a^{-1}. Of this source, 7% (10 Tg a^{-1}) is estimated to be anthropogenic SOA from urban and/or industrial VOCs, 64% (90 Tg a^{-1}) is estimated to be anthropogenically controlled SOA formed primarily from biogenic VOCs, and 9% (13 Tg a^{-1}) is estimated to be SOA of biogenic origin without anthropogenic input. SOA from biomass burning VOCs account for 3% (3 Tg a^{-1}) of the global SOA source, and oxidation of POA (mostly BBOA) accounts for 16% (23 Tg a^{-1}) of the SOA source. The scatter plots of measured vs. modeled OOA concentrations using the initial (biogenic SOA-dominated) and final optimized models are shown in Fig. 13. Most of the field measurements included in this study were obtained in the northern hemisphere. Thus, the usefulness of this approach for constraining global SOA sources will be further enhanced when more factor analysis results from the southern hemisphere become available.

In addition, the OA factor results from six AMS field measurement campaigns were used by Ervens et al. [115] to evaluate the extent to which simple assumptions of OA composition and mixing states can reproduce measured CCN number concentrations in ambient air. These authors concluded that a simple treatment of CCN composition and/or mixing state as a function of distance from sources predicts CCN concentration and cloud drop concentration with reasonably good accuracy (~15%). This finding may have important implications on large-scale modeling of aerosol–cloud interactions. Furthermore, Wex et al. [127] reported the utilization of the global OA factor results shown in Fig. 9 for evaluating global-scale CCN predictions.

Conclusions and future research

The development in recent years of quantitative aerosol mass spectrometers capable of reporting highly time-resolved organic aerosol composition data has enabled a new application of factor analysis techniques. The most useful techniques are based on mass conservation and enable estimation of the time series and mass spectra of OA factors, representing the contributions of hundreds of different chemical species, and which can be associated with different sources and/or processes. Aerodyne AMS instruments have so far dominated this field, with only three publications using data from other real-time aerosol mass spectrometers. PMF is a positively constrained, error-weighed variant of the bilinear model and is most commonly used in factor-analysis applications. The criteria for selecting the optimum PMF solutions are complex, and have been summarized here. The uncertainty of the PMF factors may be evaluated on the basis of variations in plausible solutions obtained with different FPEAK and SEED values or via bootstrapping analysis with replacement of mass spectra. Some advanced applications have been reported using, e.g., a unified dataset combining aerosol mass spectrometry and gas-phase VOC data, additional constraints on the mass spectra, or three-dimensional models. The components most commonly identified in the data include several primary emissions factors, for example hydrocarbon-like OA, biomass-burning OA, and cooking OA, and several secondary process-related factors under the umbrella of oxygenated organic aerosols (OOAs). Fresh OOAs are more variable whereas aged OOAs seem to converge into non-volatile and highly oxygenated OAs which are likely to be acid-dominated. The results from these analyses are being used to evaluate multiple models including both aerosol process and regional and global chemical-transport models.

Some methodological aspects require further research, for example improvement of error estimates, better methods for determining proper relative weights when mixing datasets, and methods to account for the variability in component mass spectra. Factor analysis of datasets acquired with improved soft ionization aerosol mass spectrometers, which retain information of individual organic molecules and larger frag-

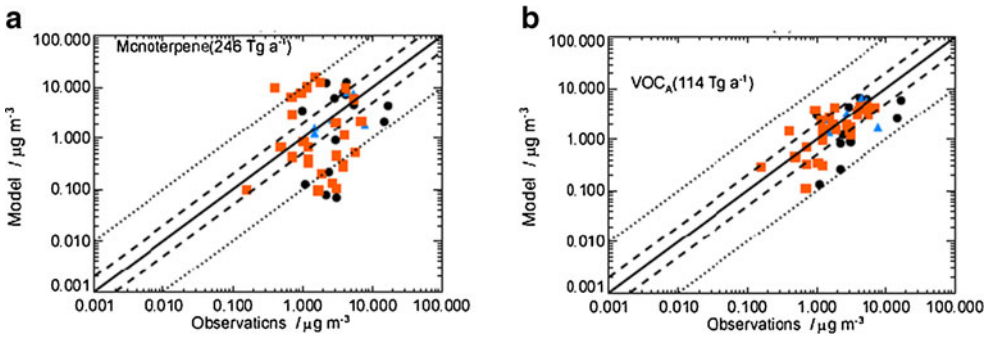

Fig. 13 Scatterplot of simulated (GLOMAP) versus observed (AMS; data shown in Fig. 9) OOA data using (**a**) the initial biogenic dominated model and (**b**) the final optimized model. Observation locations are classified as urban (*black circles*), urban-downwind (*blue triangles*), and rural and/or remote (*red squares*) as in Zhang et al. [6]. The 1:1 line (*solid*), 2:1 lines (*dashed*), and 10:1 lines (*dotted*) are shown. (Figs. 3d and 4d, respectively, in Ref. [126]. Copyright 2011, reprinted with permission)

ments, and of multidimensional datasets (e.g., acquired from thermodenuder–AMS and fast chromatography–MS) might enable further resolution of some of the current uncertainties associated with OA sources and processes.

Finally, more efforts should be placed on integrated analysis and interpretation of the factor analysis results of worldwide aerosol mass spectrometry datasets acquired across very wide geographic locations and on different time scales. These results enable the development of holistic and simplified views on OA climatology necessary for constraining global and regional models and for advancing current knowledge of the roles of aerosols in global climate change and degradation of human health and ecosystems.

Acknowledgements This study was supported by the Office of Science (BER), US Department of Energy, Atmospheric Systems Research (ASR) Program (Grants No. DE-FG02-08ER64627, DE-SC0002191, DE-FG02-11ER65293, DE-FGOZO5ER63982, and DE-SC0006035), NSF NSF ATM-0919189, and ATM-0528170, and by NOAA NA08OAR4310565, NA05OAR4310102, AB133R04SE0480, and RA1330-02-SE-015O.

References

1. IPCC (2007) In: Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds) Climate change 2007: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University, New York
2. Ghan S, Schwartz SE (2007) Aerosol properties and processes: a path from field and laboratory measurements to global climate models. Bull Am Meteorol Soc 88(7):1059–1083. doi:10.1175/BAMS-88-7-1059
3. NARSTO (2003) Particulate matter science for policy makers: a NARSTO Assessment, EPRI 1007735
4. Pope CA, Ezzati M, Dockery DW (2009) Fine-particulate air pollution and life expectancy in the United States. New Engl J Med 360:376–386. doi:10.1056/NEJMsa0805646
5. Bytnerowicz A, Fenn ME (1996) Nitrogen deposition in California forests - A review. Environ Pollut 92(2):127–146. doi:10.1016/0269-7491(95)00106-9
6. Zhang Q et al (2007) Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced northern hemisphere mid-latitudes. Geophys Res Lett 34: L13801. doi:10.1029/2007GL029979
7. Kanakidou M et al (2005) Organic aerosol and global climate modelling: a review. Atmos Chem Phys 5:1053–1123. doi:10.5194/acp-5-1053-2005
8. Seinfeld JH, Pankow JF (2003) Organic atmospheric particulate material. Annu Rev Phys Chem 54:121–140. doi:10.1146/annurev.physchem.54.011002.103756
9. Goldstein A, Galbally I (2007) Known and unexplored organic constituents in the earth's atmosphere. Environ Sci Technol 41 (5):1514–1521. doi:10.1021/es072476p
10. de Gouw J, Jimenez JL (2009) Organic aerosols in the earth's atmosphere. Environ Sci Technol 43:7614–7618. doi:10.1021/es9006004
11. Hallquist M et al (2009) The formation, properties and impact of secondary organic aerosol: current and emerging issues. Atmos Chem Phys 9:5155–5236. doi:10.5194/acp-9-5155-2009
12. Canagaratna MR et al (2010) Evolution of vehicle exhaust particles in the atmosphere. J Air Waste Manag Assoc 60 (10):1192–1203. doi:10.3155/1047-3289.60.10.1192
13. Jimenez JL et al (2009) Evolution of organic aerosols in the atmosphere. Science 326:1525–1529. doi:10.1126/science.1180353
14. Ng NL et al (2010) Organic aerosol components observed in Northern Hemispheric datasets from aerosol mass spectrometry. Atmos Chem Phys 10:4625–4641. doi:10.5194/acp-10-4625-2010
15. Blando JD, Turpin BJ (2000) Secondary organic aerosol formation in cloud and fog droplets: a literature evaluation of plausibility. Atmos Environ 34(10):1623–1632. doi:10.1016/S1352-2310(99)00392-1
16. Kroll JH, Seinfeld JH (2008) Chemistry of secondary organic aerosol: formation and evolution of low-volatility organics in the atmosphere. Atmos Environ 42:3593–3624. doi:10.1016/j.atmosenv.2008.01.003
17. Donahue NM et al (2006) Coupled partitioning, dilution, and chemical aging of semivolatile organics. Environ Sci Technol 40 (8):2635–2643. doi:10.1021/es052297c
18. Morgan WT et al (2010) Airborne measurements of the spatial distribution of aerosol chemical composition across Europe and evolution of the organic fraction. Atmos Chem Phys 10:4065–4083. doi:10.5194/acp-4010-4065-2010
19. Shrivastava MK et al (2007) Sources of organic aerosol: positive matrix factorization of molecular marker data and comparison of results from different source apportionment models. Atmos Environ 41(40):9353–9369. doi:10.1016/j.atmosenv.2007.09.016
20. Jaeckels JM, Bae MS, Schauer JJ (2007) Positive matrix factorization (PMF) analysis of molecular marker measurements to quantify the sources of organic aerosols. Environ Sci Technol 41(16):5763–5769. doi:10.1021/es062536b
21. Dutton SJ et al (2010) Source apportionment using positive matrix factorization on daily measurements of inorganic and organic speciated PM2.5. Atmos Environ 44(23):2731–2741. doi:10.1016/j.atmosenv.2010.04.038
22. Sun Y et al (2011) Characterization and source apportionment of water-soluble organic matter in atmospheric fine particles ($PM_{2.5}$) with high-resolution aerosol mass spectrometry and GC-MS. Environ Sci Technol 45(11):4854–4861. doi:10.1021/es200162h
23. Pratt KA, Prather KA (2011) Mass spectrometry of atmospheric aerosols—recent developments and applications. Part II: On-line mass spectrometry techniques. Mass Spectrometry Reviews. doi:10.1002/mas.20330
24. Williams BJ et al (2006) An in-situ instrument for speciated organic composition of atmospheric aerosols: thermal desorption aerosol GC/MS-FID (TAG). Aerosol Sci Technol 40 (8):627–638. doi:10.1080/02786820600754631
25. Canagaratna MR et al (2007) Chemical and microphysical characterization of ambient aerosols with the aerodyne aerosol mass spectrometer. Mass Spectrom Rev 26:185–222. doi:10.1002/mas.20115
26. Smith JN et al (2004) Atmospheric measurements of sub-20 nm diameter particle chemical composition by thermal desorption chemical ionization mass spectrometry. Aerosol Sci Technol 38 (2):100–110. doi:10.1080/02786820490249036
27. Dreyfus MA, Johnston MV (2008) Rapid sampling of individual organic aerosol species in ambient air with the photoionization aerosol mass spectrometer. Aerosol Sci Technol 42(1):18–27. doi:10.1080/02786820701785112

28. Murphy DM (2007) The design of single particle laser mass spectrometers. Mass Spectrom Rev 26:150–165
29. Johnston MV, Wang SY, Reinard MS (2006) Nanoparticle mass spectrometry: pushing the limit of single particle analysis. Appl Spectrosc 60(10):264A–272A
30. DeCarlo P et al (2004) Particle morphology and density characterization by combined mobility and aerodynamic diameter measurements. Part 1: theory. Aerosol Sci Technol 38:1185–1205. doi:10.1080/027868290903907
31. Aiken AC et al (2008) O/C and OM/OC ratios of primary, secondary, and ambient organic aerosols with a high resolution time-of-flight aerosol mass spectrometer. Environ Sci Technol 42(12):4478–4485. doi:10.1021/es703009q
32. DeCarlo PF et al (2006) A field-deployable high-resolution time-of-flight aerosol mass spectrometer. Anal Chem 78(24):8281–8289. doi:10.1021/ac061249n
33. Kroll JH et al (2011) Carbon oxidation state as a metric for describing the chemistry of atmospheric organic aerosol. Nat Chem 3:133–139. doi:10.1038/NCHEM.94
34. Dreyfus MA et al (2009) Organic aerosol source apportionment from highly time-resolved molecular composition measurements. Atmos Environ 43(18):2901–2910. doi:10.1016/j.atmosenv.2009.03.008
35. Öktem B, Tolocka MP, Johnston MV (2004) On-line analysis of organic components in fine and ultrafine particles by photoionization aerosol mass spectrometry. Anal Chem 76:253–261. doi:10.1021/ac0350559
36. Northway MJ et al (2007) Demonstration of a VUV lamp photoionization source for improved organic speciation in an aerosol mass spectrometer. Aerosol Sci Technol 41(9):828–839. doi:10.1080/02786820701496587
37. Zahardis J, Geddes S, Petrucci GA (2011) Improved understanding of atmospheric organic aerosols via innovations in soft ionization aerosol mass spectrometry. Anal Chem 83(7):2409–2415. doi:10.1021/ac102737k
38. Ulbrich I et al (2009) Interpretation of organic components from positive matrix factorization of aerosol mass spectrometric data. Atmos Chem Phys 9:2891–2918. doi:10.5194/acp-9-2891-2009
39. Sun Y et al (2011) Characterization of the sources and processes of organic and inorganic aerosols in New York City with a high-resolution time-of-flight aerosol mass spectrometer. Atmos Chem Phys 11:1581–1602. doi:10.5194/acp-11-1581-2011
40. Zhang Q et al (2005) Hydrocarbon-like and oxygenated organic aerosols in Pittsburgh: insights into sources and processes of organic aerosols. Atmos Chem Phys 5:3289–3311. doi:10.5194/acp-5-3289-2005
41. Zhang Q et al (2005) Deconvolution and quantification of hydrocarbon-like and oxygenated organic aerosols based on aerosol mass spectrometry. Environ Sci Technol 39(13):4938–4952. doi:10.1021/es048568l
42. Volkamer R et al (2006) Secondary organic aerosol formation from anthropogenic VOCs: Rapid and higher than expected. Geophys Res Lett 33:L17811. doi:10.1029/2006GL026899
43. Takegawa N et al (2006) Seasonal and diurnal variations of submicron organic aerosols in Tokyo observed using the Aerodyne aerosol mass spectrometer (AMS). J Geophys Res 111:D11206. doi:10.1029/2005JD006515
44. Zhang Q et al (2006) Component analysis of organic aerosols in urban, rural, and remote atmospheres based on aerosol mass spectrometry. in *7th International Aerosol Conference*. St. Paul, Minnesota
45. Bae M-S et al (2007) Interference of organic signals in highly-time resolved nitrate measurements by low mass resolution aerosol mass spectrometry. J Geophys Res 112:D22305. doi:10.1029/2007JD008614
46. Cottrell LD et al (2008) Submicron particles at Thompson Farm during ICARTT measured using aerosol mass spectrometry. J Geophys Res 113:D08212. doi:10.1029/2007JD009192
47. Ziemba LD et al (2010) Characterization of Aerosol associated with enhancement of number concentrations of small particles in a suburban forested environment. J Geophys Res 115:D12206. doi:10.1029/2009JD012614
48. Paatero P, Tapper U (1994) Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values. Environmetrics 5:111–126. doi:10.1002/env.3170050203
49. Paatero P (1997) Least squares formulation of robust non-negative factor analysis. Chemom Intell Lab Syst 37(1):23–35. doi:10.1016/S0169-7439(96)00044-5
50. Slowik JG et al (2010) Simultaneous factor analysis of organic particle and gas mass spectra: AMS and PTR-MS measurements at an urban site. Atmos Chem Phys 10:1969–1988. doi:10.5194/acp-10-1969-2010
51. Lanz VA et al (2007) Source apportionment of submicron organic aerosols at an urban site by factor analytical modelling of aerosol mass spectra. Atmos Chem Phys 7:1503–1522. doi:10.5194/acp-7-1503-2007
52. Ulbrich IM et al (2011) Three-dimensional factorization of size-resolved organic aerosol mass spectra from Mexico City. Atmos Meas Tech Discuss 4:4561–4630. doi:10.5194/amtd-4-4561-2011
53. Paatero P (1999) The multilinear engine - A table-driven, least squares program for solving multilinear problems, including the n-way parallel factor analysis model. J Comput Graph Stat 8 (4):854–888. doi:10.2307/1390831
54. Paatero P et al (2002) Understanding and controlling rotations in factor analytic models. Chemom Intell Lab Syst 60(1–2):253–264. doi:10.1016/S0169-7439(01)00200-3
55. Zhang Q et al (2005) *Time and size-resolved chemical composition of submicron particles in Pittsburgh - Implications for aerosol sources and processes*. J Geophys Res 110:D07S09. doi:10.1029/2004JD004649
56. Lanz VA et al (2010) Characterization of Aerosol Chemical Composition with aerosol mass spectrometry in Central Europe: an overview. Atmos Chem Phys 10453–10471. doi:10.5194/acp-10-10453-2010
57. Huffman JA et al (2009) Chemically-resolved volatility measurements of organic aerosol from different sources. Environ Sci Technol 43:5351–5357. doi:10.1021/es803539d
58. Huffman JA et al (2009) Chemically-resolved aerosol volatility measurements from two megacity field studies. Atmos Chem Phys 9(18):7161–7182. doi:10.5194/acp-9-7161-2009
59. Raatikainen T et al (2010) Physicochemical properties and origin of organic groups detected in boreal forest using an aerosol mass spectrometer. Atmos Chem Phys 10(4):2063–2077. doi:10.5194/acp-10-2063-2010
60. Cappa CD, Jimenez JL (2010) Quantitative estimates of the volatility of ambient organic aerosol. Atmos Chem Phys 10 (12):5409–5424. doi:10.5194/acp-10-5409-2010
61. DeCarlo PF et al (2008) Fast airborne aerosol size and chemistry measurements with the high resolution aerosol mass spectrometer during the MILAGRO Campaign. Atmos Chem Phys 8:4027–4048
62. DeCarlo PF et al (2010) Investigation of the sources and processing of organic aerosol over the Central Mexican Plateau from aircraft measurements during MILAGRO. Atmos Chem Phys 10(12):5257–5280. doi:10.5194/acp-10-5257-2010
63. Docherty KS et al (2008) Apportionment of primary and secondary organic aerosols in Southern California during the 2005 Study of Organic Aerosols in Riverside (SOAR-1). Environ Sci Technol 42:7655–7662. doi:10.1021/es8008166

64. Aiken AC et al (2009) Mexico City aerosol analysis during MILAGRO using high resolution aerosol mass spectrometry at the urban supersite (T0) – Part 1: fine particle composition and organic source apportionment. Atmos Chem Phys 9:6633–6653. doi:10.5194/acpd-9-8377-2009
65. Aiken AC et al (2010) Mexico City aerosol analysis during MILAGRO using high resolution aerosol mass spectrometry at the urban supersite (T0) – Part 2: analysis of the biomass burning contribution and the non-fossil carbon fraction. Atmos Chem Phys 10:5315–5341. doi:10.5194/acp-10-5315-2010
66. Huang X-F et al (2010) Highly time-resolved chemical characterization of atmospheric submicron particles during 2008 Beijing Olympic Games using an Aerodyne High Resolution aerosol mass spectrometer. Atmos Chem Phys 10:8933–8945. doi:10.5194/acp-10-8933-2010
67. Slowik JG et al (2010) Characterization of a large biogenic secondary organic aerosol event from eastern Canadian forests. Atmos Chem Physics 10(6):2825–2845. doi:10.5194/acp-10-2825-2010
68. Huang XF et al (2011) Characterization of submicron aerosols at a rural site in Pearl River Delta of China using an Aerodyne High-Resolution aerosol mass spectrometer. Atmos Chem Phys 11(5):1865–1877. doi:10.5194/acp-11-1865-2011
69. Williams BJ et al (2010) Major components of atmospheric organic aerosol in southern California as determined by hourly measurements of source marker compounds. Atmos Chem Phys 10(23):11577–11603. doi:10.5194/acp-10-11577-2010
70. Williams BJ et al (2007) Chemical speciation of organic aerosol during the International Consortium for Atmospheric Research on Transport and Transformation 2004: results from in situ measurements. J Geophys Res 112:D10S26. doi:10.1029/2006JD007601
71. Allan JD et al (2010) Contributions from transport, solid fuel burning and cooking to primary organic aerosols in two UK cities. Atmos Chem Phys 10:647–668. doi:10.5194/acp-10-647-2010
72. Sun J et al (2010) Highly time- and size-resolved characterization of submicron aerosol particles in Beijing using an aerodyne aerosol mass spectrometer. Atmos Environ 44(1):131–140. doi:10.1016/j.atmosenv.2009.03.20
73. Allan JD et al (2003) Quantitative sampling using an Aerodyne aerosol mass spectrometer. Part 1: techniques of data interpretation and error analysis. J Geophys Res-Atmos 108(D3):4090. doi:10.1029/2002JD002358
74. Paatero P, Hopke PK (2003) Discarding or downweighting high-noise variables in factor analytic models. Anal Chim Acta 490(1–2):277–289. doi:10.1016/S0003-2670(02)01643-4
75. Nemitz E et al (2008) An eddy-covariance system for the measurement of surface/atmosphere exchange fluxes of submicron aerosol chemical species—first application above an urban area. Aerosol Sci Technol 42:636–657. doi:10.1080/02786820802227352
76. Canagaratna MR et al (2004) Chase studies of particulate emissions from in-use New York city vehicles. Aerosol Sci Technol 38:555–573. doi:10.1080/02786820490465504
77. Bahreini R et al (2005) Measurements of secondary organic aerosol from oxidation of cycloalkanes, terpenes, and m-xylene using the Aerodyne aerosol mass spectrometer. Environ Sci Technol 39(15):5674–5688. doi:10.1021/es048061a
78. Alfarra MR et al (2006) A Mass Spectrometric study of secondary organic aerosols formed from the photooxidation of anthropogenic and biogenic precursors in a reaction chamber. Atmos Chem Phys 6:5279–5293. doi:10.5194/acp-6-5279-2006
79. Schneider J et al (2006) Mass spectrometric analysis and aerodynamic properties of various types of combustion-related aerosol particles. Int J Mass Spec 258(1–3):37–49. doi:10.1016/j.ijms.2006.07.008
80. Alfarra MR et al (2007) Identification of the mass spectral signature of organic aerosols from wood burning emissions. Environ Sci Technol 41(16):5770–5777. doi:10.1021/es062289b
81. Weimer S et al (2008) Organic aerosol mass spectral signatures from wood-burning emissions: influence of burning conditions and wood type. J Geophys Res-Atmos 113:D10304. doi:10.1029/2007JD009309
82. Kiendler-Scharr A et al (2009) Aerosol mass spectrometric features of biogenic SOA: observations from a plant chamber and in rural atmospheric environments. Environ Sci Technol 43 (21):8166–8172. doi:10.1021/es901420b
83. Mohr C et al (2009) Characterization of primary organic aerosol emissions from meat cooking, trash burning, and motor vehicles with high-resolution aerosol mass spectrometry and comparison with ambient and chamber observations. Environ Sci Technol 43 (7):2443–2449. doi:10.1021/es8011518
84. He LY et al (2010) Characterization of high-resolution aerosol mass spectra of primary organic aerosol emissions from Chinese cooking and biomass burning. Atmos Chem Phys 10(23):11535–11543. doi:10.5194/acpd-10-21237-2010
85. Lanz VA et al (2008) Source attribution of submicron organic aerosols during wintertime inversions by advanced factor analysis of aerosol mass spectra. Environ Sci Technol 42:214–220. doi:10.1021/es0707207
86. Ng NL et al (2011) An Aerosol Chemical Speciation Monitor (ACSM) for routine monitoring of the composition and mass concentrations of ambient aerosol. Aerosol Sci Technol 45 (7):770–784. doi:10.1080/02786826.2011.560211
87. Kondo Y et al (2007) Oxygenated and water-soluble organic aerosols in Tokyo. J Geophys Res 112:D01203. doi:10.1029/2006JD007056
88. Herndon SC et al (2008) The correlation of secondary organic aerosol with odd oxygen in Mexico City. Geophys Res Lett 35: L15804. doi:10.11029/12008GL034058
89. Wood EC et al (2010) Investigation of the correlation between odd oxygen and secondary organic aerosol in Mexico City and Houston. Atmos Chem Phys 10:8947–8968. doi:10.5194/acp-10-8947-2010
90. Minguillon MC et al (2011) Fossil versus contemporary sources of fine elemental and organic carbonaceous particulate matter during the DAURE campaign in Northeast Spain. Atmos Chem Phys Discuss 11:23573–23618. doi:10.5194/acpd-11-23573-2011
91. Hildebrandt L et al (2010) Aged organic aerosol in the Eastern Mediterranean: the Finokalia Aerosol Measurement Experiment-2008. Atmos Chem Phys 10(9):4167–4186. doi:10.5194/acp-10-4167-2010
92. Hawkins LN, Russell LM (2010) Oxidation of ketone groups in transported biomass burning aerosol from the 2008 Northern California Lightning Series fires. Atmos Environ 44(34):4142–4154. doi:10.1016/j.atmosenv.2010.07.036
93. Allan JD et al (2003) Quantitative sampling using an Aerodyne aerosol mass spectrometer. Part 2: measurements of fine particulate chemical composition in two UK Cities. J Geophys Res-Atmos 108(D3):4091. doi:10.1029/2002JD002359
94. Alfarra MR et al (2004) Characterization of urban and regional organic aerosols in the lower Fraser Valley using two Aerodyne aerosol mass spectrometers. Atmos Environ 38:5745–5758. doi:10.1016/j.atmosenv.2004.01.054
95. Zhang Q et al (2004) Insights into the chemistry of new particle formation and growth events in Pittsburgh based on aerosol mass spectrometry. Environ Sci Technol 38 (18):4797–4809. doi:10.1021/es035417u
96. Salcedo D et al (2006) Characterization of ambient aerosols in Mexico City during the MCMA-2003 Campaign with aerosol

mass spectrometry: results from the CENICA Supersite. Atmos Chem Phys 6:925–946. doi:10.5194/acp-6-925-2006
97. Schneider J et al (2005) Nucleation particles in diesel exhaust: composition Inferred from in situ mass spectrometric analysis. Environ Sci Technol 39(16):6153–6161. doi:10.1021/es049427m
98. Bon DM et al (2011) Measurements of volatile organic compounds at a suburban ground site (T1) in Mexico City during the MILAGRO 2006 campaign: measurement comparison, emission ratios, and source attribution. Atmos Chem Phys 11 (6):2399–2421. doi:10.5194/acp-11-2399-2011
99. Paatero P, Hopke PK (2007) Rotational tools for factor analytic models. J Chemometrics 23:91–100
100. Chow JC, Watson JG (2002) Review of PM2.5 and PM10 apportionment for fossil fuel combustion and other sources by the chemical mass balance receptor model. Energy Fuels 16 (2):222–260. doi:10.1021/ef0101715
101. Ng NL et al (2011) Real-time methods for estimating organic component mass concentrations from aerosol mass spectrometer data. Environ Sci Technol 45(3):910–916. doi:10.1021/es102951k
102. Paatero P (1997) A weighted non-negative least squares algorithm for three-way 'PARAFAC' factor analysis. Chemom Intell Lab Syst 38(2):223–242. doi:10.1016/S0169-7439(97)00031-2
103. Ng NL et al (2011) Changes in organic aerosol composition with aging inferred from aerosol mass spectra. Atmos Chem Phys 11:6465–6474. doi:10.5194/acp-11-6465-2011
104. Russell LM, Bahadur R, Ziemann PJ (2011) Identifying organic aerosol sources by comparing functional group composition in chamber and atmospheric particles. Proc Natl Acad Sci U S A 108(9):3516–3521. doi:10.1073/pnas.1006461108
105. Durant JL et al (2010) Short-term variation in near-highway air pollutant gradients on a winter morning. Atmos Chem Phys 10 (17):8341–8352. doi:10.5194/acp-10-8341-2010
106. Cross ES et al (2009) Single particle characterization using a light scattering module coupled to a time-of-flight aerosol mass spectrometer. Atmos Chem Phys 9(20):7769–7793. doi:10.5194/acp-9-7769-2009
107. Huffman JA et al (2008) Development and characterization of a fast-stepping/scanning thermodenuder for chemically-resolved aerosol volatility measurements. Aerosol Sci Technol 42(5):395–407. doi:10.1080/02786820802104981
108. Sun Y et al (2009) Size-resolved aerosol chemistry on whistler mountain, Canada with a high-resolution aerosol mass spectrometer during INTEX-B. Atmos Chem Phys 9:3095–3111. doi:10.5194/acp-9-3095-2009
109. Alfarra MR (2004) Insights into atmospheric organic aerosols using an aerosol mass spectrometer, in chemical engineering. University of Manchester, Manchester
110. Duplissy J et al (2011) Relating hygroscopicity and composition of organic aerosol particulate matter. Atmos Chem Phys 11 (3):1155–1165. doi:10.5194/acp-11-1155-2011
111. Donahue NM et al (2011) A two-dimensional volatility basis set: 1. organic-aerosol mixing thermodynamics. Atmos Chem Phys 11:3303–3318. doi:10.5194/acp-11-3303-2011
112. Cubison MJ et al (2006) The characterisation of pollution aerosol in a changing photochemical environment. Atmos Chem Phys 6:5573–5588. doi:10.5194/acp-6-5573-2006
113. Cubison MJ et al (2008) The influence of chemical composition and mixing state of Los Angeles urban aerosol on CCN number and cloud properties. Atmos Chem Phys 8:5649–5667. doi:10.5194/acp-8-5649-2008
114. Quinn PK et al (2008) Influence of particle size and chemistry on the cloud nucleating properties of aerosols. Atmos Chem Phys 8 (4):1029–1042. doi:10.5194/acp-8-1029-2008
115. Ervens B et al (2010) CCN predictions using simplified assumptions of organic aerosol composition and mixing state: a synthesis from six different locations. Atmos Chem Phys 10:4795–4807. doi:10.5194/acp-10-4795-2010
116. Wang J et al (2010) The importance of aerosol mixing state and size-resolved composition on CCN concentration and the variation of the importance with atmospheric aging of aerosols. Atmos Chem Phys 10(15):7267–7283. doi:10.5194/acpd-10-11751-2010
117. Duplissy J et al (2008) Cloud forming potential of secondary organic aerosol under near atmospheric conditions. Geophys Res Lett 35:L03818. doi:10.1029/2007GL031075
118. Massoli P et al (2010) Relationship between aerosol oxidation level and hygroscopic properties of laboratory generated secondary organic aerosol (SOA) particles. Geophys Res Lett 37:L24801. doi:10.21029/22010GL045258
119. Robinson AL et al (2007) Rethinking organic aerosols: semivolatile emissions and photochemical aging. Science 315:1259. doi:10.1126/science.1133061
120. Hodzic A et al (2010) Modeling organic aerosols in a megacity: potential contribution of semi-volatile and intermediate volatility primary organic compounds to secondary organic aerosol formation. Atmos Chem Phys 10:5491–5514. doi:10.5194/acp-10-5491-2010
121. Dzepina K et al (2009) Evaluation of recently-proposed secondary organic aerosol (SOA) models for a case study in Mexico City. Atmos Chem Phys 9:5681–5709. doi:10.5194/acp-9-5681-2009
122. Dzepina K et al (2011) Modeling the multiday evolution and aging of secondary organic aerosol during MILAGRO 2006. Environ Sci Technol 45(8):3496–3503. doi:10.1021/es103186f
123. Fast J et al (2009) Evaluating simulated primary anthropogenic and biomass burning organic aerosols during MILAGRO: implications for assessing treatments of secondary organic aerosols. Atmos Chem Phys 9(16):6191–6215. doi:10.5194/acp-9-6191-2009
124. Tsimpidi AP et al (2010) Evaluation of the volatility basis-set approach for the simulation of organic aerosol formation in the Mexico City metropolitan area. Atmos Chem Phys 10(2):525–546. doi:10.5194/acp-10-525-2010
125. Grieshop AP, Donahue NM, Robinson AL (2009) Laboratory investigation of photochemical oxidation of organic aerosol from wood fires 2: analysis of aerosol mass spectrometer data. Atmos Chem Phys 9(6):2227–2240. doi:10.5194/acp-9-2227-2009
126. Spracklen DV et al (2011) Measurement-based constraints on the global secondary organic aerosol budget. Atmos Chem Phys Discuss 11:5699–5755. doi:10.5194/acpd-11-5699-2011
127. Wex H et al (2010) The influence of the external mixing state of atmospheric aerosol on derived CCN number concentrations. Geophys Res Lett 37:L10805. doi:10.1029/2010GL043337
128. Dzepina K, Arey J, Marr L, Worsnop D, Salcedo D, Zhang Q, Onasch TB, Molina L, Molina M, Jimenez J (2007) Detection of particle-phase Polycyclic Aromatic Hydrocarbons (PAHs) in Mexico City using an Aerosol Mass Spectrometer. Int J Mass Spectrom 263:152–170